



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 99742

TO: Minh-Tam Davis

Location:

Art Unit: 1642

July 25, 2003

8E12

Case Serial Number: 09/991681

From: P. Sheppard

Location: CM1-1E03

Phone: (703) 308-4499

sheppard@uspto.gov

Search Notes

THIS PAGE BLANK (USPTO)

99699

From: Chan, Christina
Sent: Friday, July 25, 2003 10:19 AM
To: Davis, Minh-Tam; STIC-Biotech/ChemLib
Subject: RE: Rush search request for 09/991681

RECEIVED

JUL 25 2003

(STIC)

Please rush. Thanks Chris

Chris Chan

TC 1600 New Hire Training Coordinator and SPE 1644
308-3973
CM-1, 9B19

-----Original Message-----

From: Davis, Minh-Tam
Sent: Friday, July 25, 2003 10:18 AM
To: Chan, Christina
Subject: Rush search request for 09/991681

Please search for interference only:
SEQ IS NO:27-31.

Thank you.

MINH TAM DAVIS

ART UNIT 1642, ROOM 8A01, MB 8E12
305-2008

Edward Hart
Technical Info. Specialist
STIC/Biotech
CMI 6B02 Tel: 305-9203

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 7/25/03
Searcher Prep/Review: 7/28/03
Clerical: _____
Online time: _____

TYPE OF SEARCH:

NA Sequences: _____
AA Sequences: 5
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)

STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: 92
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

Sheppard, Paula

919742

From: Davis, Minh-Tam
Sent: Tuesday, July 22, 2003 1:10 PM
To: Sheppard, Paula
Subject: Rush search request for 09/991681

I am sorry that this is the third time that I ask you for this case. Since I could not find any decent art even for fragments of SEQ ID NO:27, could you please search in commercial database, PG PUB and issued patent files:

Oligomer search for SEQ ID NOs: 27-31.

Thank you.

If you want me to submit a new search request to the library, please let me know.

MINH TAM DAVIS

ART UNIT 1642, ROOM 8A01, MB 8E12
305-2008

Point of Contact
P. Sheppard
Telephone number: (703) 308-4499

7/25/03

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM: protein - protein search, using sw model

Run on: July 24, 2003, 11:58:19 / Search time 12.4832 Seconds
(without alignments)
623.047 Million cell updates/sec

Title: US-09-991-681-31

Perfect score: 49
Sequence: 1 EPLGPRGDSPLRPFQHL.....SAGPELLRQDKRPPSGSTGS 49

Scoring table: OLIGO
Gapop 60.0, Gapext 60.0

Searched: 1107863 seqs, 158726573 residues

Word size: 0

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

A: Geneseq_19Jun03:*

1: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA2001.DAT:*
23: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA2002.DAT:*
24: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	100.0	49	23	AAO19169 Human prostate-spe
2	49	100.0	49	23	AAO19169 Human prostate-spe
3	49	100.0	518	19	AAO19169 Human prostate-spe
4	49	100.0	518	23	AAO19169 Human prostate-spe
5	49	100.0	518	23	AAO19169 Human prostate-spe
6	49	100.0	1807	22	ABG09728 Human prostate-spe
7	49	100.0	1839	23	ABP64835 Human prostate-spe
8	49	100.0	1982	22	ABG09731 Human prostate-spe
9	8	16.3	32	22	ABG49938 Human liver peptid

10	8	16.3	32	22	ABR29911
11	8	16.3	32	22	ABR35089
12	8	16.3	32	22	ABR35089
13	8	16.3	32	22	ABR35089
14	8	16.3	32	22	ABR35089
15	8	16.3	32	22	ABR35089
16	8	16.3	32	22	ABR35089
17	8	16.3	32	22	ABR35089
18	8	16.3	32	22	ABR35089
19	8	16.3	32	22	ABR35089
20	7	14.3	101	22	ABG39938
21	7	14.3	115	22	AAU87586
22	7	14.3	115	22	AAU87586
23	7	14.3	164	22	AAU87294
24	7	14.3	234	21	AAU87294
25	7	14.3	280	22	AAU87294
26	7	14.3	314	21	AAU87294
27	7	14.3	314	21	AAU87294
28	7	14.3	401	22	AAU87294
29	7	14.3	418	22	AAU87294
30	7	14.3	418	22	AAU87294
31	7	14.3	418	22	AAU87294
32	7	14.3	421	22	AAU87294
33	7	14.3	461	22	AAU87294
34	7	14.3	499	23	ABU03835
35	7	14.3	529	24	ABU03835
36	7	14.3	564	24	ABU03835
37	7	14.3	612	22	ABU03835
38	7	14.3	977	22	ABU03835
39	7	14.3	989	24	ABU03835
40	7	14.3	1006	22	ABU03835
41	7	14.3	1162	16	AAU87294
42	7	14.3	1519	21	AAU87294
43	6	12.2	10	22	AAU87294
44	6	12.2	12	21	AAU87294
45	6	12.2	13	14	AAU87294

ALIGNMENTS

RESULT 1	AAO19169	AAO19169 standard; Protein; 49 AA.
ID	AAO19169	AAO19169 standard; Protein; 49 AA.
AC	AAO19169	AAO19169 standard; Protein; 49 AA.
XX	AAO19169	AAO19169 standard; Protein; 49 AA.
DT	27-NOV-2002	(first entry)
DE	Human prostate-specific PS18 protein fragment #5.	
XX	Human, prostate; prostate-specific sequence; prostate cancer; PS18;	
KW	EST; expressed sequence tag; cytosolic; gene therapy.	
XX	Human, prostate; prostate-specific sequence; prostate cancer; PS18;	
OS	Homo sapiens.	
XX	US2002086316-A1.	
XX	04-JUL-2002.	
PD	26-NOV-2001; 2001US-0991681.	
XX	23-APR-1998; 98US-0065383.	
XX	23-APR-1997; 97US-0842385.	
PR	23-APR-1997; 97US-0842385.	
XX	(BIL)/ BILINGEL P A.	
PA	(COHE)/ COHEN M.	
PA	(COLP)/ COLPITS T L.	
PA	(FRIE)/ FRIEDMAN P N.	
PA	(GORD)/ GORDAN J.	
PA	(GRAN)/ GRANADOS E N.	
PA	(HODG)/ HODGES S C.	
PA	(KLAS)/ KLAS M R.	

Peptide #2562 enco
Peptide #2595 enco
Protein #2506 enco
Human brain expres
Human bone marrow
Peptide #2537 enco
Peptide #2631 enco
Peptide #2512 enco
Human peptide enco
Novel human diagno
Novel human diagno
Novel central nerv
Novel signal trans
Novel central nerv
Human secreted pro
Human polypeptide
Human colon cancer
Human colon cancer
Human protein kina
Human cell signal
Human protein sequ
Amino acid sequenc
Human polypeptide
Human polypeptide
Human drug metabol
Human expressed pr
Human adipocyte se
Drosophila melanog
Drosophila melanog
Human transporter
Human stromalin-2.
Human tyrosine kin
Human complementar
Anti-IL12 antibody
Mouse OSE-2 antige

PA (KRAT/) KRATOCHVIL J D.
 PA (ROBE/) ROBERTS-RAPP L.
 PA (RUSS/) RUSSELL J C.
 PA (STRO/) STROUPE S D.
 XX
 PI Billingel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;
 PI Granados EN, Hodges SC, Klass MR, Kratochvil JD, Roberts-Rapp L;
 PI Russell JC, Stroupe SD;
 XX
 DR WPI; 2002-665429/71.
 XX
 PT Novel PS18 polypeptide for detecting, diagnosing, staging, monitoring,
 PT prognosticating, preventing, treating, or determining predisposition of
 PT individual to diseases and conditions of prostate, e.g. prostate
 PT cancer -
 XX
 PS Claim 17; Page 44; 58pp; English.
 XX
 CC The present invention relates to a number of prostate-specific sequences
 CC derived from the human PS18 gene. These can be used in the detection,
 CC monitoring and treatment of prostate diseases, particularly prostate
 CC cancer. The present sequence is a PS18 protein fragment of the
 CC invention. The coding sequences of the invention were isolated from a
 CC prostate tissue expressed sequence tag (EST) library.
 CC
 SQ Sequence 49 AA;
 XX
 Query Match 100.0%; Score 49; DB 23; Length 49;
 Best Local Similarity 100.0%; Pred. No. 3.3e-41;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 1 EPLGPRGQDSPILQRPQHMDQGMKHSFSAPELLRQDKRRSGSTGS 49
 Db 1 EPLGPRGQDSPILQRPQHMDQGMKHSFSAPELLRQDKRRSGSTGS 49
 XX
 RESULT 2
 ID AAM50813 standard; Protein; 49 AA.
 AC AAM50813;
 XX
 DT 01-MAY-2002 (first entry)
 XX
 DE PS18 prostate marker immunogenic polypeptide.
 XX
 KW PS18; prostate; marker; prostate cancer; tumour; metastasis;
 KW benign prostatic hyperplasia; prostatic intraepithelial neoplasia;
 KW prostatic; human; diagnosis; therapy; vaccine; immunogen.
 XX
 OS Homo sapiens.
 XX
 PN US2001055758-A1.
 XX
 PD 27-DEC-2001.
 XX
 PF 23-APR-1998; 98US-0065383.
 XX
 PR 23-APR-1997; 97US-0842385.
 XX
 PA (BILL/) BILLING-MEDEL P A.
 PA (COHE/) COHEN M.
 PA (COPL/) COPLITTIS T L.
 PA (FRIE/) FRIEDMAN P N.
 PA (GORD/) GORDON J.
 PA (GRAN/) GRANADOS E N.
 PA (HODG/) HODGES S C.
 PA (KLAS/) KLAS M R.
 PA (KRAT/) KRATOCHVIL J D.
 PA (ROBE/) ROBERTS-RAPP L.
 XX
 PI Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;
 PI Granados EN, Hodges SC, Klass MR, Kratochvil JD, Roberts-Rapp L;
 PI Russell JC, Stroupe SD;
 XX

PI Russell JC, Stroupe SD;
 XX
 DR WPI; 2002-187683/24.
 XX
 PT Detecting presence of target PS18 polynucleotide in test sample,
 PT useful for detecting, diagnosing, staging, monitoring, prognosticating,
 PT preventing or treating or determining predisposition to prostate
 PT disease -
 XX
 PS Claim 17; Page 44; 57pp; English.
 XX
 CC The present sequence is that of an immunogenic polypeptide
 CC comprising amino acids 393-441 of human prostate-specific PS18
 CC polypeptide (see AAM50809). A PS18 consensus sequence (see
 CC ABA91651) is found at least 12 times more often in prostate than
 CC in non-prostate tissue. PS18 polypeptides, polynucleotides,
 CC antibodies, agonists and inhibitors are useful for detecting,
 CC diagnosing, staging, monitoring, prognosticating, preventing and
 CC treating (including by genetic immunisation), or determining the
 CC predisposition of an individual to, diseases and conditions of the
 CC prostate, such as benign prostatic hyperplasia, prostaticis,
 CC prostatic intraepithelial neoplasia, prostate cancer, tumours and
 CC metastases. The PS18 polypeptides can be produced by expression
 CC of PS18 polynucleotides in transfected host cells, and
 CC immunogenic peptides are useful for raising PS18-specific
 CC antibodies of diagnostic use. The methods and reagents of the
 CC invention may provide an early means of detecting diseases of the
 CC prostate and may also provide new markers which can differentiate
 CC between the clinically important and unimportant prostate cancers
 CC without the use of surgery.
 CC
 SQ Sequence 49 AA;
 XX
 Query Match 100.0%; Score 49; DB 23; Length 49;
 Best Local Similarity 100.0%; Pred. No. 3.3e-41;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 1 EPLGPRGQDSPILQRPQHMDQGMKHSFSAPELLRQDKRRSGSTGS 49
 Db 1 EPLGPRGQDSPILQRPQHMDQGMKHSFSAPELLRQDKRRSGSTGS 49
 XX
 RESULT 3
 ID AAM85472 standard; Protein; 518 AA.
 AC AAM85472;
 XX
 DT 25-MAR-2003 (updated)
 DT 25-FEB-1999 (first entry)
 XX
 DE PS18 protein encoded by consensus sequence.
 XX
 KW EST clone; PS18; prostate tumour tissue; prostatic disease; cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO9848054-A1.
 XX
 PD 29-OCT-1998.
 XX
 PF 23-APR-1998; 98WO-US08239.
 XX
 PR 23-APR-1997; 97US-0842385.
 XX
 PA (ABBO) ABBOTT LAB.
 XX
 PI Billing-medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;
 PI Granados EN, Hodges SC, Klass MR, Kratochvil JD, Robertstrapp L;
 PI Russell JC, Stroupe SD;
 XX
 DR WPI; 1998-610000/51.
 DR N-PSDB; AAV82812.
 XX

XX New P118 nucleic acid and proteins - used for diagnosis and
PT treatment of prostatic disease, especially cancer, and also for drug
PT screening
XX
PS Claim 17; Page 93-94; 117pp; English.
XX
PS The present sequence is encoded by consensus P118 sequence derived from
CC overlapping EST (expressed sequence tag) clones AAV82803-10. The clones
CC were identified from a cDNA library made from prostate tumour tissue.
CC Recombinant P118 protein is used to detect P118-specific antibodies,
CC to raise antibodies for detection of P118 antigens, to screen for
CC specific binding agents (potential therapeutics), and to isolate specific
CC antibodies from serum. Detection of P118 protein or nucleic acid, which
CC are prostate related, and altered or elevated in prostatic disease, is
CC used for detection, diagnosis, staging, monitoring and prognosis of
CC prostatic disease, particularly cancer, and to identify subjects at
CC risk.
CC (Updated on 25-MAR-2003 to correct PI field.)
CC
SQ Sequence 518 AA;
XX
Query Match 100.0%; Score 49; DB 19; Length 518;
Best Local Similarity 100.0%; Pred. No. 2.4e-40;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
OY 1 EPLGRGDSPLLRPOHLMDOGQWRHSFSAPELLRQDKPRSGSTGS 49
DB 393 EPLGRGDSPLLRPOHLMDOGQWRHSFSAPELLRQDKPRSGSTGS 441
XX
RESULT 4
ID AAO19165 standard; Protein; 518 AA.
XX
AC AAO19165;
XX
DT 27-NOV-2002 (first entry)
XX
DE Human prostate-specific P118 protein fragment #1.
XX
KW Human; prostate; prostate-specific sequence; prostate cancer; P118;
KW EST; expressed sequence tag; cytostatic; gene therapy.
XX
OS Homo sapiens.
XX
PN US2002086316-A1.
XX
PD 04-JUL-2002.
XX
PF 26-NOV-2001; 2001US-0991681.
XX
PR 23-APR-1998; 98US-0065383.
PR 23-APR-1997; 97US-0842385.
XX
XX (BILL/) BILLINGEL P A.
PA (COHE/) COHEN M.
PA (COLP/) COLPITTS T L.
PA (FRIE/) FRIEDMAN P N.
PA (GORD/) GORDAN J.
PA (GRAN/) GRANADOS E N.
PA (HODG/) HODGES S C.
PA (KLAS/) KLAS M R.
PA (KRAT/) KRATOCHVIL J D.
PA (ROBE/) ROBERTS-RAPP L.
PA (RUSSE/) RUSSELL J C.
PA (STROU/) STROUPE S D.
XX
PI Billingsel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;
PI Granados EN, Hodges SC, Kلاس MR, Kratochvil JD, Roberts-Rapp L;
PI Russell JC, Stroupe SD;
XX
DR WPI; 2002-665429/71.

XX Novel P118 polypeptide for detecting, diagnosing, staging, monitoring,
PT prognosticating, preventing, treating, or determining predisposition of
PT individual to diseases and conditions of prostate, e.g. prostate
PT cancer -
XX
PS Claim 17; Page 42-43; 58pp; English.
XX
PS The present invention relates to a number of prostate-specific sequences
CC derived from the human P118 gene. These can be used in the detection,
CC monitoring and treatment of prostate diseases, particularly prostate
CC cancer. The present sequence is a P118 protein fragment of the
CC invention. The coding sequence of the invention were isolated from a
CC prostate tissue expressed sequence tag (EST) library.
XX
SQ Sequence 518 AA;
XX
Query Match 100.0%; Score 49; DB 23; Length 518;
Best Local Similarity 100.0%; Pred. No. 2.4e-40;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
OY 1 EPLGRGDSPLLRPOHLMDOGQWRHSFSAPELLRQDKPRSGSTGS 49
DB 393 EPLGRGDSPLLRPOHLMDOGQWRHSFSAPELLRQDKPRSGSTGS 441
XX
RESULT 5
ID AAM50809 standard; Protein; 518 AA.
XX
AC AAM50809;
XX
DT 01-MAY-2002 (first entry)
XX
DE P118 prostate marker partial sequence.
XX
KW P118; prostate; marker; prostate cancer; tumour; metastasis;
KW benign prostatic hyperplasia; prostatic intraepithelial neoplasia;
KW prostatitis; human; diagnosis; therapy; vaccine.
XX
OS Homo sapiens.
XX
PN US2001055758-A1.
XX
PD 27-DEC-2001.
XX
PF 23-APR-1998; 98US-0065383.
XX
PR 23-APR-1997; 97US-0842385.
XX
XX (BILL/) BILLING-MEDEL P A.
PA (COHE/) COHEN M.
PA (COLP/) COLPITTS T L.
PA (FRIE/) FRIEDMAN P N.
PA (GORD/) GORDAN J.
PA (GRAN/) GRANADOS E N.
PA (HODG/) HODGES S C.
PA (KLAS/) KLAS M R.
PA (KRAT/) KRATOCHVIL J D.
PA (ROBE/) ROBERTS-RAPP L.
XX
PI Billings-Medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;
PI Granados EN, Hodges SC, Kلاس MR, Kratochvil JD, Roberts-Rapp L;
PI Russell JC, Stroupe SD;
XX
DR WPI; 2002-187683/24.
DR N-PSDB; ABA91651.
XX
XX Detecting presence of target P118 polynucleotide in test sample,
PT useful for detecting, diagnosing, staging, monitoring, prognosticating,
PT preventing or treating or determining predisposition to prostate
PT disease -
XX

PS Claim 17, Page 42-43; 57pp; English.

XX The present sequence is that of a human prostate-specific PS118
 CC polypeptide, as predicted from a partial consensus cDNA sequence
 CC (see ABA91651), and lacking the N-terminal region. The PS118
 CC consensus sequence is found at least 12 times more often in
 CC prostate than in non-prostate tissue. PS118 polypeptides,
 CC including derivatives of the present sequence, polynucleotides,
 CC antibodies, agonists and inhibitors are useful for detecting,
 CC diagnosing, staging, monitoring, prognosticating, preventing and
 CC treating (including by genetic immunisation), or determining the
 CC predisposition of an individual to, diseases and conditions of the
 CC prostate, such as benign prostatic hyperplasia, prostaticitis, of the
 CC prostatic, interepithelial neoplasia, prostate cancer, tumours and
 CC metastases. The PS118 polypeptides can be produced by expression
 CC of PS118 polynucleotides in transfected host cells. The methods
 CC and reagents of the invention may provide an early means of
 CC detecting diseases of the prostate and may also provide new markers
 CC which can differentiate between the clinically important and
 CC unimportant prostate cancers without the use of surgery.

XX SQ Sequence 518 AA;

Query Match 100.0%; Score 49; DB 23; Length 518;
 Best Local Similarity 100.0%; Pred. No. 2.4e-40;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EPLGPRGQDSPLLQRPQHLMDQGQMRHSFSAGPELLRQDKRPRSSTGS 49
 |||
 DB 393 EPLGPRGQDSPLLQRPQHLMDQGQMRHSFSAGPELLRQDKRPRSSTGS 441

RESULT 6
 ID ABO9728 standard; Protein; 1807 AA.

XX ABO9728;
 AC
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #9719.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 PI Drmanac RT, Liu C, Tang YT;
 PI
 XX
 DR WPI: 2001-639362/73.
 DR N-PSDB; AAS73915.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX
 PS Claim 20; SEQ ID No 40087; 103pp; English.
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABO0010-ABG3077 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 1807 AA;

Query Match 100.0%; Score 49; DB 22; Length 1807;
 Best Local Similarity 100.0%; Pred. No. 6.7e-40;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EPLGPRGQDSPLLQRPQHLMDQGQMRHSFSAGPELLRQDKRPRSSTGS 49
 |||
 DB 1682 EPLGPRGQDSPLLQRPQHLMDQGQMRHSFSAGPELLRQDKRPRSSTGS 1730

RESULT 7
 ID ABP64835 standard; Protein; 1839 AA.

XX ABP64835;
 AC
 XX
 DT 25-FEB-2003 (first entry)
 XX
 DE Human protein SEQ ID 495.
 XX
 KW Human; expressed sequence tag; EST;
 KW haematopoietic disorder; central nervous system disease; viral infection;
 KW peripheral nervous system disease; non-healing wound; infectious disease;
 KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;
 KW fungal infection; autoimmune disorder; coagulation disorder; noctropic;
 KW antiallergic; antiinflammatory; immunosuppressive; neuroprotective;
 KW cytosolic; haemostatic; virucide; antibacterial; fungicide;
 KW immunostimulant; cerebroprotective.
 XX
 OS Homo sapiens.
 XX
 PN WO200259260-A2.
 XX
 PD 01-AUG-2002.
 XX
 PF 16-NOV-2001; 2001WO-US42950.
 XX
 PR 17-NOV-2000; 2000US-0714936.
 XX
 PA (HYSE-) HYSEQ INC.
 PI Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;
 PI Ren F, Xue Ao, Yang Y, Wehrman T, Drmanac RT;
 PI
 XX
 DR WPI: 2002-590824/63.
 DR N-PSDB; ABO99421.
 XX
 PT New isolated polynucleotide, useful in research, diagnostic or
 PT therapeutic methods, e.g. preventing or treating disorders involving
 PT aberrant protein expression or biological activity -
 XX
 PS Claim 20; SEQ ID 495; 394pp; English.

CC The present invention relates to novel human coding sequences
 CC (ABG99268-ABG99608) and proteins (ABP64682-ABP65022). The sequences are
 CC useful in therapeutic, diagnostic and research methods. The
 CC polynucleotides may be used in the field of molecular biology as
 CC hybridisation probes, primers for PCR, for chromosome and gene mapping,
 CC for the recombinant production of protein, or in generation of anti-sense
 CC DNA or RNA. The polynucleotides are useful in diagnostics as expressed
 CC sequence tags (ESTs) for identifying expressed genes or for physical
 CC mapping of the human genome. The proteins may be used as molecular weight
 CC markers, or as nutritional sources or supplements. The proteins may be
 CC used to maintain and expand cell population in a totipotent or
 CC pluripotent state useful for re-engineering damaged or diseased
 CC tissues, transplantation, manufacture of bio-pharmaceuticals or the
 CC development of bio-sensors. The polynucleotides and proteins are useful
 CC for preventing, treating or ameliorating disorders involving aberrant
 CC protein expression or biological activity, e.g. haematopoietic disorders,
 CC central/peripheral nervous system diseases, mechanical and traumatic
 CC disorders, non-healing wounds, immune deficiencies and disorders,
 CC infectious diseases caused by viral, bacterial or fungal infection,
 CC autoimmune disorders, allergic reactions and conditions, coagulation
 CC disorders, or cancer. The polynucleotide sequences of the invention were
 CC assembled from ESTs isolated mainly by sequencing by hybridisation, and
 CC in some cases, sequences obtained from one or more public databases.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 1839 AA;

Query Match 100.0%; Score 49; DB 23; Length 1839;
 Best Local Similarity 100.0%; Pred. No. 6.8e-40;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EPLGPRGDSPLLRPOHMDQGMHSFSAPELLRDKRPRSGSTGS 49
 DB 1714 EPLGPRGDSPLLRPOHMDQGMHSFSAPELLRDKRPRSGSTGS 1762

RESULT 8
 ABG09731

ID ABG09731 standard; Protein; 1982 AA.

AC ABG09731;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #9722.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PE 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

PI Dmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

DR N-PSDB; AAS733918.

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -

XX Claim 20; SEQ ID No 40090; 103bp; English.

PS The invention relates to isolated polynucleotide (I) and
 XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping; and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 1982 AA;

Query Match 100.0%; Score 49; DB 22; Length 1982;
 Best Local Similarity 100.0%; Pred. No. 7.3e-40;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EPLGPRGDSPLLRPOHMDQGMHSFSAPELLRDKRPRSGSTGS 49
 DB 1796 EPLGPRGDSPLLRPOHMDQGMHSFSAPELLRDKRPRSGSTGS 1844

RESULT 9
 ABG49938

ID ABG49938 standard; Peptide; 32 AA.

AC ABG49938;

DT 25-FEB-2003 (first entry)

DE Human liver peptide, SEQ ID No 28586.

KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
 KW hypercholesterolaemia; coronary heart disease.

OS Homo sapiens.

PN WO200157273-A2.

PD 09-AUG-2001.

PE 30-JAN-2001; 2001WO-US00664.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-488898/53.

PT Human genome-derived single exon nucleic acid probes useful for
 PT analysing gene expression in human adult liver -

XX Claim 27; SEQ ID No 28586; 658bp; English.
PS
XX The invention relates to a single exon nucleic acid probe (SENP) (1) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult
CC liver. (1) may be used for predicting, measuring and displaying gene
CC expression in samples derived from human adult liver. The genes
CC identified may be involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which
CC is associated with coronary heart disease. ABG47348-ABG59930 represent
CC human liver single exon encoded peptides of the invention.
CC Note: The sequence information for this patent does not appear in the
CC printed specification but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 32 AA;
Query Match 16.3%; Score 8; DB 22; Length 32;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 27 HSFSAGPE 34
Db 13 HSFSAGPE 20
RESULT 10
ABB29911
ID ABB29911 standard; Peptide; 32 AA.
AC ABB29911;
XX 01-FEB-2002 (first entry)
DT
XX Peptide #2562 encoded by breast cell single exon nucleic acid probe.
DE
XX Human; microarray; single exon probe; gene expression; breast;
XX disease; cancer.
XX Homo sapiens.
OS
XX MO200157271-A2.
PN
XX 09-AUG-2001.
PD
XX 30-JAN-2001; 2001WO-US00662.
PF
XX 04-FEB-2000; 2000US-0180312.
PR
XX 26-MAY-2000; 2000US-0207456.
PR
XX 30-JUN-2000; 2000US-0608408.
PR
XX 03-AUG-2000; 2000US-0632366.
PR
XX 21-SEP-2000; 2000US-0234687.
PR
XX 27-SEP-2000; 2000US-0236359.
PR
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-496933/54.
XX
XX WPI; 2001-496933/54.
XX
XX New spatially-addressable set of single exon nucleic acid probes,
XX useful for measuring gene expression in sample derived from human
XX breast, comprises number of single exon nucleic acid probes -
XX
XX Claim 27; SEQ ID NO 12879; 327bp + sequence listing; English.
XX
XX The invention relates to a spatially-addressable set of single exon
XX nucleic acid probes for measuring gene expression in a sample derived
XX from human breast and BT 474 cells. The method involves contacting

CC the probes with a collection of detectably labelled nucleic acids
CC derived from mRNA of human breast, and then measuring the label
CC bound to each probe of the microarray. The probes are useful for
CC verifying the expression of regions of genomic DNA predicted to
CC encode proteins. They are useful for gene discovery, and for
CC determining predisposition and/or prognosing breast disease. Gene
CC expression analysis is useful for assessing the toxicity of chemical
CC agents on cells. The microarray of this invention presents a far greater
CC diversity of probes for measuring gene expression, with far less bias
CC than expressed sequence tag microarrays. The method is suitable for
CC rapid production of functional information from genomic sequence. The
CC present sequence is a peptide encoded by a single exon nucleic acid
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 32 AA;
Query Match 16.3%; Score 8; DB 22; Length 32;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 27 HSFSAGPE 34
Db 13 HSFSAGPE 20
RESULT 11
ABB35089
ID ABB35089 standard; Peptide; 32 AA.
AC ABB35089;
XX 04-FEB-2002 (first entry)
DT
XX Peptide #2595 encoded by human foetal liver single exon probe.
DE
XX Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
XX Homo sapiens.
OS
XX MO200157271-A2.
PN
XX 09-AUG-2001.
PD
XX 30-JAN-2001; 2001WO-US00669.
PF
XX 04-FEB-2000; 2000US-0180312.
PR
XX 26-MAY-2000; 2000US-0207456.
PR
XX 30-JUN-2000; 2000US-0608408.
PR
XX 03-AUG-2000; 2000US-0632366.
PR
XX 21-SEP-2000; 2000US-0234687.
PR
XX 27-SEP-2000; 2000US-0236359.
PR
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483447/52.
XX
XX WPI; 2001-483447/52.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human foetal liver -
XX
XX Claim 27; SEQ ID NO 27724; 639bp + sequence listing; English.
XX
XX The invention relates to a single exon nucleic acid probe for
XX measuring human gene expression in a sample derived from human foetal
XX liver. The single exon nucleic acid probes may be used for predicting,
XX measuring and displaying gene expression in samples derived from human
XX foetal liver. The present sequence is a peptide encoded by a single exon
XX nucleic acid probe of the invention.

CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 32 AA;

Query Match 16.3%; Score 8; DB 22; Length 32;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 HSFSGAPE 34
DB 13 HSFSGAPE 20

RESULT 12
ABR20507
ID ABR20507 standard; Protein; 32 AA.

AC ABR20507;

DT 23-JAN-2002 (first entry)

DE Protein #2506 encoded by probe for measuring heart cell gene expression.

KM Human; gene expression; heart; microarray; vascular system;

KW cardiovascular disease; hypertension; cardiac arrhythmia;

KX congenital heart disease.

OS Homo sapiens.

PN WO200157274-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00666.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-488899/53.

PT Single exon nucleic acid probes for analyzing gene expression in human

PT hearts -

PS Claim 15; SEQ ID No 22277; 530pp; English.

XX The present invention relates to single exon nucleic acid probes for

XX measuring human gene expression in a sample derived from human heart (see

XX ABR21535-ABR41305). The present sequence is a protein encoded by one such

XX probe. The probes may be used for predicting, measuring and displaying

XX gene expression in samples derived from the human heart via microarrays.

XX By measuring gene expression, the probes are useful for predicting,

XX diagnosing, grading, staging, monitoring and prognosing diseases of the

XX human heart and vascular system e.g. cardiovascular disease,

XX hypertension, cardiac arrhythmias and congenital heart disease.

XX Note: The sequence data for this patent did not form part of the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 32 AA;

XX Query Match 16.3%; Score 8; DB 22; Length 32;

XX Best Local Similarity 100.0%; Pred. No. 1;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 HSFSGAPE 34
DB 13 HSFSGAPE 20

RESULT 13
AAM55910
ID AAM55910 standard; Protein; 32 AA.

AC AAM55910;

DT 05-NOV-2001 (first entry)

DE Human brain expressed single exon probe encoded protein SEQ ID NO: 28015.

KM Human; brain expressed exon; gene expression analysis; probe;

KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;

KX epilepsy; cancer.

OS Homo sapiens.

PN WO200157275-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00667.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-483446/52.

PT Single exon nucleic acid probes for analyzing gene expression in human

PT brains -

PS Example 4; SEQ ID NO: 28015; 650pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid

XX probes which are derived from genomic sequences expressed in the human

XX brain. They can be used to measure gene expression in brain cell samples,

XX which may enable the diagnosis and improved treatment of nervous system

XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,

XX epilepsy and cancers. The present sequence is a protein encoded by one of

XX the probes of the invention.

XX Sequence 32 AA;

XX Query Match 16.3%; Score 8; DB 22; Length 32;

XX Best Local Similarity 100.0%; Pred. No. 1;

XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX QY 27 HSFSGAPE 34

XX DB 13 HSFSGAPE 20

XX RESULT 14

XX AAM68280

XX ID AAM68280 standard; Protein; 32 AA.

XX AC AAM68280;

XX XX

```

DT 06-NOV-2001 (first entry)
XX Human bone marrow expressed probe encoded protein SEQ ID NO: 28586.
DE Human bone marrow expressed exon; gene expression analysis; probe;
XX microarray; cancer; leukemia; lymphoma; myeloma.
XX Homo sapiens.
XX WO200157276-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US00668.
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488900/53.
XX DR WPI; 2001-488900/53.
XX PT Human genome-derived single exon nucleic acid probes useful for
XX PT analyzing gene expression in human bone marrow -
XX PS Example 4; SEQ ID NO: 28586; 658bp + Sequence Listing; English.
XX CC The present invention provides a number of single exon nucleic acid
XX CC probes which are derived from genomic sequences expressed in the human
XX CC bone marrow. They can be used to measure gene expression in bone marrow
XX CC samples, which may enable the improved diagnosis and treatment of cancers
XX CC such as lymphoma, leukemia and myeloma. The present sequence is a
XX CC protein encoded by one of the probes of the invention.
XX SQ Sequence 32 AA;

Query Match 16.3%; Score 8; DB 22; Length 32;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 HSFSAGPE 34
DB 13 HSFSAGPE 20

RESULT 15
AAM16103
ID AAM16103 standard; Protein; 32 AA.
XX AAM16103;
AC AAM16103;
XX 12-OCT-2001 (first entry)
XX Peptide #2537 encoded by probe for measuring cervical gene expression.
DE Peptide #2537 encoded by probe for measuring cervical gene expression.
XX Probe; human; microarray; gene expression; cervical epithelial cell;
XX KW cervical cancer.
XX OS Homo sapiens.
XX WO200157276-A2.
XX PN 09-AUG-2001.
XX PD 30-JAN-2001; 2001WO-US00670.
XX PF
XX

```

```

PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488901/53.
XX DR WPI; 2001-488901/53.
XX PT Human genome-derived single exon nucleic acid probes useful for
XX PT analyzing gene expression in human cervical epithelial cells -
XX PS Claim 27; SEQ ID No 20929; 487bp; English.
XX CC The present invention relates to human single exon nucleic acid probes
XX CC (SENP: see A1110068-A128459). The present sequence is a peptide encoded
XX CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs
XX CC can be used to produce a single exon microarray, which can be used for
XX CC measuring human gene expression in a sample derived from human cervical
XX CC epithelial cells. By measuring gene expression, the probes are therefore
XX CC useful in grading and/or staging of diseases of the cervix, notably
XX CC cervical cancer.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 32 AA;

Query Match 16.3%; Score 8; DB 22; Length 32;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 HSFSAGPE 34
DB 13 HSFSAGPE 20

```

Search completed: July 24, 2003, 12:21:21
Job time : 15.4832 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 12:18:29 ; Search time 5.66764 Seconds
(without alignments)
365.801 Million cell updates/sec

Title: US-09-991-681-31

Perfect score: 49
Sequence: 1 EPLGPRGQDSPLQRPQHL.....SAGPELLRQDKRPRSGSTGS 49

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310858 residues

Word size : 0

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Database :

Issued Patents, AA: *
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep: *
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep: *
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep: *
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep: *
5: /cgn2_6/ptodata/1/1aa/PCTUS.COMB.pep: *
6: /cgn2_6/ptodata/1/1aa/Backfltest1.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query length	ID	Description
1	49	100.0	49	US-09-065-383-31	Sequence 31, Appl
2	49	100.0	518	US-09-065-383-27	Sequence 27, Appl
3	7	14.3	173	US-09-252-991A-17084	Sequence 17084, A
4	7	14.3	285	US-09-107-532A-6376	Sequence 6376, Ap
5	7	14.3	729	US-09-252-991A-32535	Sequence 32535, A
6	6	12.2	13	US-08-426-627-10	Sequence 10, Appl
7	6	12.2	114	US-09-203-623-2	Sequence 2, Appl
8	6	12.2	114	US-09-523-462-2	Sequence 2, Appl
9	6	12.2	114	US-09-522-980-2	Sequence 2, Appl
10	6	12.2	139	US-09-252-991A-32472	Sequence 32472, A
11	6	12.2	153	US-09-252-991A-24763	Sequence 24763, A
12	6	12.2	171	US-09-252-991A-29708	Sequence 29708, A
13	6	12.2	205	US-09-252-991A-29378	Sequence 29378, A
14	6	12.2	226	US-09-489-847-215	Sequence 215, App
15	6	12.2	247	US-09-252-991A-32129	Sequence 32129, A
16	6	12.2	249	US-09-252-991A-21188	Sequence 21188, A
17	6	12.2	261	US-09-489-847-371	Sequence 371, App
18	6	12.2	272	US-09-107-532A-5898	Sequence 5898, Ap
19	6	12.2	282	US-09-252-991A-17674	Sequence 17674, A
20	6	12.2	290	US-09-585-858-49	Sequence 49, Appl
21	6	12.2	303	US-08-985-950-2	Sequence 2, Appl
22	6	12.2	303	US-09-546-049-2	Sequence 2, Appl
23	6	12.2	309	US-09-252-991A-17975	Sequence 17975, A
24	6	12.2	325	US-09-252-991A-31554	Sequence 31554, A
25	6	12.2	331	US-09-134-001C-5188	Sequence 5188, Ap
26	6	12.2	365	US-09-252-991A-21955	Sequence 21955, A
27	6	12.2	399	US-09-252-991A-26424	Sequence 26424, A

28	6	12.2	401	US-09-252-991A-20438	Sequence 20438, A
29	6	12.2	413	US-09-252-991A-26239	Sequence 26239, A
30	6	12.2	450	US-09-191-879-2	Sequence 2, Appl
31	6	12.2	450	US-09-252-991A-32284	Sequence 32284, A
32	6	12.2	454	US-09-470-512A-8	Sequence 8, Appl
33	6	12.2	456	US-09-252-991A-28231	Sequence 28231, A
34	6	12.2	461	US-09-107-532A-4308	Sequence 4308, Ap
35	6	12.2	466	US-09-038-909-2	Sequence 2, Appl
36	6	12.2	466	US-08-868-435-2	Sequence 2, Appl
37	6	12.2	466	US-08-744-231-2	Sequence 2, Appl
38	6	12.2	466	US-09-470-512A-9	Sequence 9, Appl
39	6	12.2	466	US-09-543-141-2	Sequence 2, Appl
40	6	12.2	466	US-09-636-499-7	Sequence 7, Appl
41	6	12.2	466	US-09-273-871A-12	Sequence 12, Appl
42	6	12.2	467	US-09-134-001C-5168	Sequence 5168, Ap
43	6	12.2	474	US-09-252-991A-27558	Sequence 27558, A
44	6	12.2	504	US-09-252-991A-28279	Sequence 28279, A
45	6	12.2	574	US-09-252-991A-29351	Sequence 29351, A

ALIGNMENTS

RESULT 1
US-09-065-383-31
Sequence 31, Application US/09065383
Patent No. 6391543
GENERAL INFORMATION:
APPLICANT: BILTING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GORDON, JULIAN
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KLAS, MICHAEL R.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: ROBERTS-RAPP, LISA
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FaestSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/065,383
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/842,385
FILING DATE: 23-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6084, US, P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6391543e
US-09-065-383-31

Query Match 100.0%; Score 49; DB 4; Length 49;
Best Local Similarity 100.0%; Pred. No. 5.6e-41;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EPLGRGDSPLLRPOHMDQGMHSPFAGPELLRQDKRPRSGTGS 49
Db 1 EPLGRGDSPLLRPOHMDQGMHSPFAGPELLRQDKRPRSGTGS 49

RESULT 2

US-09-065-383-27
Sequence 27, Application US/09065383
Patent No. 6391543

GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GORDON, JULIAN
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KLAS, MICHAEL R.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: ROBERTS-RAP, LISA
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
TITLE OF INVENTION: FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/065,383
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/842,385
FILING DATE: 23-APR-1997

ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6084.US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:

INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 518 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: No. 6391543e
US-09-065-383-27

Query Match 100.0%; Score 49; DB 4; Length 518;
Best Local Similarity 100.0%; Pred. No. 4.1e-40;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EPLGRGDSPLLRPOHMDQGMHSPFAGPELLRQDKRPRSGTGS 49
Db 393 EPLGRGDSPLLRPOHMDQGMHSPFAGPELLRQDKRPRSGTGS 441

RESULT 3

US-09-252-991A-17084
Sequence 17084, Application US/09252991A
Patent No. 6551795

GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 17084
LENGTH: 173
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17084

Query Match 14.3%; Score 7; DB 4; Length 173;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 42 PRSGSTG 48
Db 88 PRSGSTG 94

RESULT 4

US-09-107-532A-6376
Sequence 6376, Application US/09107532A
Patent No. 6583275

GENERAL INFORMATION:
APPLICANT: Lynn A. Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354

COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998

APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 893-5007

TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 6376:
SEQUENCE CHARACTERISTICS:
LENGTH: 285 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc_feature
LOCATION: (8) LOCATION 1...285
SEQUENCE DESCRIPTION: SEQ ID NO: 6376:
US-09-107-532A-6376

Query Match 14.3%; Score 7; DB 4; Length 285;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 29 FSAGPEL 35
Db 234 FSAGPEL 240

RESULT 5
US-09-252-991A-32535
Sequence 32535, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 32535
LENGTH: 729
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32535

Query Match 14.3%; Score 7; DB 4; Length 729;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 RSGSTGS 49
Db 714 RSGSTGS 720

RESULT 6
US-08-426-627-10
Sequence 10, Application US/08426627
Patent No. 5756664
GENERAL INFORMATION:
APPLICANT: Amann, Egon
APPLICANT: Ootawara-Hamamoto, Yoko
APPLICANT: Kikuno, Reiko
APPLICANT: Takehita, Sunao
APPLICANT: Tezuka, Kenichi
TITLE OF INVENTION: No. 5756664e1 Protein with Bone Formation
TITLE OF INVENTION: Ability and Process for Its Production.
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
Dunner
STREET: 1300 I Street, N.W.

CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/426,627
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/036,841
FILING DATE: 25-MAR-1993
APPLICATION NUMBER: JP 4-71501
FILING DATE: 27-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Hammond, Alan W.
REGISTRATION NUMBER: 35,178
REFERENCE/DOCKET NUMBER: 02481-1285-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4400
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Mus musculus
US-08-426-627-10

Query Match 12.2%; Score 6; DB 1; Length 13;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 36 LRQDKR 41
Db 8 LRQDKR 13

RESULT 7
US-09-203-623-2
Sequence 2, Application US/09203623
Patent No. 6140084
GENERAL INFORMATION:
APPLICANT: Sheppard, Paul O.
APPLICANT: Delsher, Theresa A.
TITLE OF INVENTION: HUMAN THYROID PROTEIN, ZS1G45
FILE REFERENCE: 97-62
CURRENT APPLICATION NUMBER: US/09/203,623
CURRENT FILING DATE: 1998-12-01
EARLIER APPLICATION NUMBER: US 60/067,293
EARLIER FILING DATE: 1997-12-03
NUMBER OF SEQ ID NOS: 39
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 114
TYPE: PRT
ORGANISM: Homo sapiens
US-09-203-623-2

Query Match 12.2%; Score 6; DB 3; Length 114;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 RSGSTG 48
Db 19 RSGSTG 24

```
RESULT 8
US-09-523-462-2
; Sequence 2, Application US/09523462
; Patent No. 6486304
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: HUMAN THYROID PROTEIN, ZSIG45
; FILE REFERENCE: 97-62
; CURRENT APPLICATION NUMBER: US/09/523,462
; CURRENT FILING DATE: 2000-03-10
; EARLIER APPLICATION NUMBER: US 60/067,293
; EARLIER FILING DATE: 1997-12-03
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq For Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 114
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-523-462-2

Query Match
Best Local Similarity 12.2%; Score 6; DB 4; Length 114;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 RSGSTG 48
Db 19 RSGSTG 24

RESULT 9
US-09-522-980-2
; Sequence 2, Application US/09522980
; Patent No. 6500925
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: HUMAN THYROID PROTEIN, ZSIG45
; FILE REFERENCE: 97-62
; CURRENT APPLICATION NUMBER: US/09/522,980
; CURRENT FILING DATE: 2000-03-10
; EARLIER APPLICATION NUMBER: US 60/067,293
; EARLIER FILING DATE: 1997-12-03
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq For Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 114
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-522-980-2

Query Match
Best Local Similarity 12.2%; Score 6; DB 4; Length 114;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 RSGSTG 48
Db 19 RSGSTG 24

RESULT 10
US-09-252-991A-32472
; Sequence 32472, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1999-02-18
```

```
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32472
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32472

Query Match
Best Local Similarity 12.2%; Score 6; DB 4; Length 139;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 PRSGST 47
Db 108 PRSGST 113

RESULT 11
US-09-252-991A-24783
; Sequence 24783, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24783
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24783

Query Match
Best Local Similarity 12.2%; Score 6; DB 4; Length 153;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 RPRSGS 46
Db 9 RPRSGS 14

RESULT 12
US-09-252-991A-29708
; Sequence 29708, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29708
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29708

Query Match
Best Local Similarity 12.2%; Score 6; DB 4; Length 171;
```


Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 SGTGS 49
|||||
DB 82 SGTGS 87

RESULT 13

US-09-252-991A-29378
; Sequence 29378, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29378
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29378

Query Match 12.2%; Score 6; DB 4; Length 205;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 RPRSGS 46
|||||
DB 133 RPRSGS 138

RESULT 14

US-09-489-847-215
; Sequence 215, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031p1
; CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 215
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1226)
; OTHER INFORMATION: Xaa equals stop translation
US-09-489-847-215

Query Match 12.2%; Score 6; DB 4; Length 226;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GPRGOD 9
|||||
DB 213 GPRGOD 218

RESULT 15

US-09-252-991A-32129
; Sequence 32129, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32129
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32129

Query Match 12.2%; Score 6; DB 4; Length 247;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 RPRSGS 46
|||||
DB 145 RPRSGS 150

Search completed: July 24, 2003, 12:28:40
Job time : 7.66764 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 12:26:04 ; Search time 8.32211 Seconds
(without alignments)
699.249 Million cell updates/sec

Title: US-09-991-681-31

Perfect score: 49
Sequence: 1 EPLGPRQDSPLQRPQHLM.....SAGPELLRQDKRPRSGRTGS 49

Scoring table: OLIGO
Gapco 60.0 , Gapext 60.0

Searched: 451899 seqs, 118759770 residues

Word size : 0

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Published Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PTCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/CTCT_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16.3	32	9	US-09-864-761-35805	Sequence 35805, A
2	14.3	115	10	US-09-764-868-815	Sequence 815, App
3	14.3	135	9	US-09-916-790-31	Sequence 31, App1
4	14.3	314	9	US-09-925-299-934	Sequence 934, App
5	14.3	314	11	US-09-925-299-934	Sequence 934, App
6	14.3	314	15	US-10-106-688-4725	Sequence 4725, App
7	12.2	10	11	US-09-572-404B-104	Sequence 104, App
8	12.2	20	14	US-10-057-505-30	Sequence 30, App1
9	12.2	21	9	US-09-864-761-41759	Sequence 41759, A
10	12.2	22	14	US-10-057-505-31	Sequence 31, App1
11	12.2	27	15	US-10-057-789-260	Sequence 260, App
12	12.2	27	15	US-10-212-628-260	Sequence 260, App
13	12.2	46	9	US-09-864-761-41721	Sequence 41721, A
14	12.2	59	11	US-09-764-872-354	Sequence 354, App
15	12.2	75	9	US-09-864-761-36328	Sequence 36328, A

16	12.2	82	11	US-09-782-974C-26	Sequence 26, App1
17	12.2	100	14	US-10-016-157A-187	Sequence 187, App
18	12.2	102	14	US-10-004-281-5	Sequence 5, App1
19	12.2	114	9	US-09-850-887-1	Sequence 1, App1
20	12.2	119	9	US-09-764-870-429	Sequence 429, App
21	12.2	119	15	US-10-125-540-429	Sequence 429, App
22	12.2	131	15	US-10-102-806-482	Sequence 482, App
23	12.2	142	9	US-09-864-761-39423	Sequence 39423, A
24	12.2	149	11	US-09-764-891-4162	Sequence 4162, App
25	12.2	160	9	US-09-976-451-9	Sequence 9, App1
26	12.2	161	10	US-09-925-297-658	Sequence 658, App
27	12.2	165	10	US-09-764-864-1267	Sequence 1267, App
28	12.2	176	9	US-09-976-451-8	Sequence 8, App1
29	12.2	183	9	US-09-764-451-2	Sequence 2, App1
30	12.2	205	11	US-09-769-787-46	Sequence 46, App1
31	12.2	257	15	US-10-106-698-5221	Sequence 5221, App
32	12.2	261	9	US-09-765-205-14	Sequence 14, App1
33	12.2	276	9	US-09-244-984-6	Sequence 6, App1
34	12.2	276	15	US-10-057-321-6	Sequence 6, App1
35	12.2	290	15	US-10-270-875-49	Sequence 49, App1
36	12.2	290	15	US-10-270-878-49	Sequence 49, App1
37	12.2	290	15	US-10-270-786-49	Sequence 49, App1
38	12.2	290	15	US-10-270-710-49	Sequence 49, App1
39	12.2	290	15	US-10-270-859-49	Sequence 49, App1
40	12.2	290	16	US-10-270-846-49	Sequence 49, App1
41	12.2	297	16	US-10-080-170-86	Sequence 86, App1
42	12.2	300	9	US-09-815-242-5263	Sequence 5263, App
43	12.2	303	11	US-09-774-381-58	Sequence 58, App1
44	12.2	303	15	US-10-290-631-2	Sequence 2, App1
45	12.2	311	15	US-10-156-761-11458	Sequence 11458, A

ALIGNMENTS

RESULT 1
US-09-864-761-35805
; Sequence 35805, Application US/09864761
; Patent No. US20020048763A1
GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wenheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US/09/864, 761
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263, 6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30

```

; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 35805
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007688.15
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 35
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 31
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.9
; OTHER INFORMATION: EST_HUMAN HIT: BE531168.1, EVALUE 3.90e-01
US-09-864-761-35805

```

```

Query Match      16.3%; Score 8; DB 9; Length 32;
Best Local Similarity 100.0%; Pred. No. 0.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 27 HSFSAPE 34
    |||||
DB 13 HSFSAPE 20

```

```

RESULT 2
US-09-764-868-815
; Sequence 815, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 815
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-868-815

```

```

Query Match      14.3%; Score 7; DB 10; Length 115;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 33 PELLROD 39
    |||||
DB 77 PELLROD 83

```

```

,RESULT 3
US-09-916-790-31
; Sequence 31, Application US/09916790

```

```

; Patent No. US20020061573A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachael
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Silos-Santiago, Immaculada
; TITLE OF INVENTION: 18431 AND 32374, NOVEL HUMAN PROTEIN
; TITLE OF INVENTION: KINASE FAMILY MEMBERS AND USES THEREFOR
; FILE REFERENCE: 381552002700
; CURRENT APPLICATION NUMBER: US/09/916,790
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 60/221,543
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus amino acid sequence
US-09-916-790-31

```

```

Query Match      14.3%; Score 7; DB 9; Length 135;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 33 PELLROD 39
    |||||
DB 93 PELLROD 99

```

```

RESULT 4
US-09-925-299-934
; Sequence 934, Application US/09925299
; Patent No. US2002005627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 934
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-299-934

```

```

Query Match      14.3%; Score 7; DB 9; Length 314;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 35 LLRODKR 41
    |||||
DB 93 LLRODKR 99

```

```

RESULT 5
US-09-925-299-934
; Sequence 934, Application US/09925299
; Publication No. US20030040617A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883

```

;; PRIOR FILING DATE: 2000-03-08
;; PRIOR APPLICATION NUMBER: 60/124,270
;; PRIOR FILING DATE: 1999-03-12
;; NUMBER OF SEQ ID NOS: 1556
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 934
;; LENGTH: 314
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-925-299-934

Query Match 14.3%; Score 7; DB 11; Length 314;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 LLRODKR 41
Db 93 LLRODKR 99

RESULT 6
US-10-106-698-4725
;; Sequence 4725, Application US/10106698
;; Publication No. US20030109690A1
;; GENERAL INFORMATION:
;; APPLICANT: Ruben et al.
;; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
;; FILE REFERENCE: PA005P1
;; CURRENT APPLICATION NUMBER: US/10/106,698
;; CURRENT FILING DATE: 2002-03-27
;; PRIOR APPLICATION NUMBER: PCT/US00/26524
;; PRIOR FILING DATE: 2000-09-28
;; PRIOR APPLICATION NUMBER: US 60/157,137
;; PRIOR FILING DATE: 1999-09-29
;; PRIOR APPLICATION NUMBER: US 60/163,280
;; PRIOR FILING DATE: 1999-11-03
;; NUMBER OF SEQ ID NOS: 8564
;; SOFTWARE: PatentIn Ver. 3.0
;; SEQ ID NO 4725
;; LENGTH: 314
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-106-698-4725

Query Match 14.3%; Score 7; DB 15; Length 314;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 LLRODKR 41
Db 93 LLRODKR 99

RESULT 7
US-09-572-404B-104
;; Sequence 104, Application US/09572404B
;; Publication No. US20030078374A1
;; GENERAL INFORMATION:
;; APPLICANT: Proteom Ltd
;; TITLE OF INVENTION: Complementary peptide ligands from the human genome
;; FILE REFERENCE: Human patent
;; CURRENT APPLICATION NUMBER: US/09/572,404B
;; CURRENT FILING DATE: 2000-05-17
;; NUMBER OF SEQ ID NOS: 4203
;; SOFTWARE: ProtPatent version 1.0
;; SEQ ID NO 104
;; LENGTH: 10
;; TYPE: PRT
;; ORGANISM: Homo Sapiens
;; FEATURE:
;; OTHER INFORMATION: sequence located in BTF3L3 at 25-34 and may interact with Sequenc
US-09-572-404B-104

Query Match 12.2%; Score 6; DB 11; Length 10;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 PRSGST 47
Db 1 PRSGST 6

RESULT 8
US-10-057-505-30
;; Sequence 30, Application US/10057505
;; Publication No. US20020164674A1
;; GENERAL INFORMATION:
;; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
;; APPLICANT: AURORA BIOSCIENCES CORPORATION
;; APPLICANT: TSIEN, Roger
;; APPLICANT: HEIM, Roger
;; APPLICANT: CURBITT, Andrew
;; TITLE OF INVENTION: TANDEM FLUORESCENT PROTEIN CONSTRUCTS
;; FILE REFERENCE: REGEN1260-3
;; CURRENT APPLICATION NUMBER: US/10/057,505
;; CURRENT FILING DATE: 2002-01-25
;; PRIOR APPLICATION NUMBER: US 08/792,553
;; PRIOR FILING DATE: 1997-01-31
;; PRIOR APPLICATION NUMBER: US 09/396,003
;; PRIOR FILING DATE: 1999-09-13
;; NUMBER OF SEQ ID NOS: 31
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 30
;; LENGTH: 20
;; TYPE: PRT
;; ORGANISM: Artificial sequence
;; FEATURE:
;; OTHER INFORMATION: linker sequence
US-10-057-505-30

Query Match 12.2%; Score 6; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 SGSTGS 49
Db 12 SGSTGS 17

RESULT 9
US-09-864-761-41759
;; Sequence 41759, Application US/09864761
;; Patent No. US20020048763A1
;; GENERAL INFORMATION:
;; APPLICANT: Penn, Sharron G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wenheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; FILE REFERENCE: Aeomica-X-1
;; CURRENT APPLICATION NUMBER: US/09/864,761
;; CURRENT FILING DATE: 2001-05-23
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: GB 24263,6
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30

```

; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 41759
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AF233390.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.4
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.4
; US-09-864-761-41759

Query Match
Best Local Similarity 12.2%; Score 6; DB 9; Length 21;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 SGSTS 49
Db 8 SGSTS 13

RESULT 10
US-10-057-505-31
; Sequence 31, Application US/10057505
; Publication No. US20020164674A1
; GENERAL INFORMATION:
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: AURORA BIOSCIENCES CORPORATION
; APPLICANT: TS'EN, Roger
; APPLICANT: HEIM, Roger
; APPLICANT: CUBITT, Andrew
; TITLE OF INVENTION: TANDEM FLUORESCENT PROTEIN CONSTRUCTS
; FILE REFERENCE: REGEN1260-3
; CURRENT APPLICATION NUMBER: US/10/057,505
; TITLE OF INVENTION: ANALYSIS OF COMPLEX PROTEIN MIXTURES
; FILE REFERENCE: NADII.022CP1
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: US 08/792,553
; PRIOR FILING DATE: 1997-01-31
; PRIOR APPLICATION NUMBER: US 09/396,003
; PRIOR FILING DATE: 1999-09-13
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31
; LENGTH: 22

```

```

; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Linker sequence
; US-10-057-505-31

Query Match
Best Local Similarity 12.2%; Score 6; DB 14; Length 22;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 SGSTS 49
Db 12 SGSTS 17

RESULT 11
US-10-057-789-260
; Sequence 260, Application US/10057789
; Publication No. US20030082522A1
; GENERAL INFORMATION:
; APPLICANT: Paul Haynes
; APPLICANT: Jing Wei
; APPLICANT: John Yates
; APPLICANT: Nancy Andon
; TITLE OF INVENTION: DIFFERENTIAL LABELING FOR QUANTITATIVE
; FILE REFERENCE: NADII.022A
; CURRENT APPLICATION NUMBER: US/10/057,789
; PRIOR FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: US 60/264,576
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 60/305,232
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 311
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 260
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 8
; OTHER INFORMATION: Xaa = Modified Cysteine
; US-10-057-789-260

Query Match
Best Local Similarity 12.2%; Score 6; DB 15; Length 27;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 SGSTS 49
Db 20 SGSTS 25

RESULT 12
US-10-212-628-260
; Sequence 260, Application US/10212628
; Publication No. US20030087329A1
; GENERAL INFORMATION:
; APPLICANT: Paul Haynes
; APPLICANT: Jing Wei
; APPLICANT: John Yates
; APPLICANT: Nancy Andon
; TITLE OF INVENTION: DIFFERENTIAL LABELING FOR QUANTITATIVE
; FILE REFERENCE: NADII.022CP1
; CURRENT APPLICATION NUMBER: US/10/212,628
; PRIOR FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: US 60/264,576
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 60/305,232
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 10/057,789

```

;; PRIOR FILING DATE: 2002-01-25
;; NUMBER OF SEQ ID NOS: 311
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 260
;; LENGTH: 27
;; TYPE: PRT
;; ORGANISM: Saccharomyces cerevisiae
;; FEATURE:
;; NAME/KEY: VARIANT
;; LOCATION: 8
;; OTHER INFORMATION: Xaa = Modified Cysteine
US-10-212-628-260

Query Match 12.2%; Score 6; DB 15; Length 27;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 44 SGTGS 49
Db 20 SGTGS 25

RESULT 13
US-09-864-761-41721
; Sequence 41721, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117

;; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 41721
;; LENGTH: 46
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC015971.2
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.3
;; OTHER INFORMATION: EST HUMAN HIT: BF310134.1, EVALUATE 2.00e-15
;; OTHER INFORMATION: SWISSPROT HIT: Q56815, EVALUATE 2.80e+00
US-09-864-761-41721

Query Match 12.2%; Score 6; DB 9; Length 46;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 28 SFSAGP 33
Db 1 SFSAGP 6

RESULT 14
US-09-764-872-354
; Sequence 354, Application US/09764872
; Publication No. US20030050231A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA125
; CURRENT APPLICATION NUMBER: US/09/764,872
; PRIOR FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 957
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 354
; LENGTH: 59
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (14)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-872-354

Query Match 12.2%; Score 6; DB 11; Length 59;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 28 SFSAGP 33
Db 47 SFSAGP 52

RESULT 15
US-09-864-761-36328
; Sequence 36328, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 36328
LENGTH: 75
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL03538.1
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.2
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.5
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.1
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.4
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.9
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.4
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 11
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.6
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.6
OTHER INFORMATION: EST_HUMAN HIT: BE735129.1, EVALU6 6.00e-01
US-09-864-761-36328

Query Match 12.2% Score 6; DB 9; Length 75;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 LLORPQ 17
|||||
DB 60 LLORPQ 65

Search completed: July 24, 2003, 12:48:30
Job time : 9.32211 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 12:21:29 ; Search time 73.8228 Seconds
(without alignments)
577.701 Million cell updates/sec

Title: US-09-991-681-31

Perfect score: 49
Sequence: 1 EPLGPRGQDSPULGRPHLM.....SAGPELLRQDKPRSGSTGS 49

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 5580241 seqs, 870357830 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5580241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Pending Patents AA Main:*

1:	/cgn2_6/ptodata/1/paa/US081_COMB.pdp.*
2:	/cgn2_6/ptodata/1/paa/US082_COMB.pdp.*
3:	/cgn2_6/ptodata/1/paa/US083_COMB.pdp.*
4:	/cgn2_6/ptodata/1/paa/US084_COMB.pdp.*
5:	/cgn2_6/ptodata/1/paa/US085_COMB.pdp.*
6:	/cgn2_6/ptodata/1/paa/US086_COMB.pdp.*
7:	/cgn2_6/ptodata/1/paa/US087_COMB.pdp.*
8:	/cgn2_6/ptodata/1/paa/US088_COMB.pdp.*
9:	/cgn2_6/ptodata/1/paa/US089_COMB.pdp.*
10:	/cgn2_6/ptodata/1/paa/US090_COMB.pdp.*
11:	/cgn2_6/ptodata/1/paa/US091_COMB.pdp.*
12:	/cgn2_6/ptodata/1/paa/US092_COMB.pdp.*
13:	/cgn2_6/ptodata/1/paa/US093_COMB.pdp.*
14:	/cgn2_6/ptodata/1/paa/US094_COMB.pdp.*
15:	/cgn2_6/ptodata/1/paa/US095_COMB.pdp.*
16:	/cgn2_6/ptodata/1/paa/US096_COMB.pdp.*
17:	/cgn2_6/ptodata/1/paa/US097_COMB.pdp.*
18:	/cgn2_6/ptodata/1/paa/US098_COMB.pdp.*
19:	/cgn2_6/ptodata/1/paa/US099_COMB.pdp.*
20:	/cgn2_6/ptodata/1/paa/US100_COMB.pdp.*
21:	/cgn2_6/ptodata/1/paa/US101_COMB.pdp.*
22:	/cgn2_6/ptodata/1/paa/US102_COMB.pdp.*
23:	/cgn2_6/ptodata/1/paa/US103_COMB.pdp.*
24:	/cgn2_6/ptodata/1/paa/US104_COMB.pdp.*
25:	/cgn2_6/ptodata/1/paa/US105_COMB.pdp.*
26:	/cgn2_6/ptodata/1/paa/US106_COMB.pdp.*
27:	/cgn2_6/ptodata/1/paa/US107_COMB.pdp.*
28:	/cgn2_6/ptodata/1/paa/US108_COMB.pdp.*
29:	/cgn2_6/ptodata/1/paa/US109_COMB.pdp.*
30:	/cgn2_6/ptodata/1/paa/US110_COMB.pdp.*
31:	/cgn2_6/ptodata/1/paa/US111_COMB.pdp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49	100.0	49	12	US-08-842-385-10
2	49	100.0	49	25	US-09-991-681-31

3	49	100.0	467	12	US-08-842-385-6	Sequence 6, Appl
4	49	100.0	518	25	US-09-991-681-27	Sequence 27, Appl
5	49	100.0	1770	1	PCT-US03-01943-44	Sequence 44, Appl
6	49	100.0	1770	1	PCT-US03-01943-44	Sequence 44, Appl
7	49	100.0	1807	1	PCT-US01-08631-40087	Sequence 40087, A
8	49	100.0	1839	30	US-10-416-993-495	Sequence 495, App
9	49	100.0	1872	1	PCT-US03-04508-32	Sequence 32, Appl
10	49	100.0	1982	1	PCT-US01-08631-40090	Sequence 40090, A
11	49	100.0	2221	27	PCT-US03-01943-30	Sequence 30, Appl
12	49	100.0	2221	27	PCT-US03-01943-30	Sequence 30, Appl
13	49	100.0	2221	27	PCT-US01-08631-40087	Sequence 28863, A
14	49	100.0	32	23	US-09-864-761-35863	Sequence 35863, A
15	49	100.0	32	23	US-09-864-761-35863	Sequence 28015, A
16	49	100.0	32	27	US-10-182-993-28015	Sequence 22277, A
17	49	100.0	32	27	US-10-182-993-28015	Sequence 20929, A
18	49	100.0	32	27	US-10-182-993-28015	Sequence 12570, A
19	49	100.0	32	28	US-10-182-998-12570	Sequence 28586, A
20	49	100.0	32	28	US-10-203-134-28586	Sequence 27486, A
21	49	100.0	32	28	US-10-203-134-28586	Sequence 28586, A
22	49	100.0	32	28	US-10-203-134-28586	Sequence 28863, A
23	49	100.0	32	28	US-10-203-134-28586	Sequence 12879, A
24	49	100.0	32	28	US-10-203-134-28586	Sequence 12879, A
25	49	100.0	32	28	US-10-203-134-28586	Sequence 12879, A
26	49	100.0	32	31	US-09-724-059-17560	Sequence 17560, A
27	49	100.0	32	31	US-09-724-059-17560	Sequence 1354743, A
28	49	100.0	32	31	US-09-724-059-17560	Sequence 1354743, A
29	49	100.0	32	31	US-09-724-059-17560	Sequence 1354743, A
30	49	100.0	32	31	US-09-724-059-17560	Sequence 1354743, A
31	49	100.0	32	31	US-09-724-059-17560	Sequence 1354743, A
32	49	100.0	32	31	US-09-724-059-17560	Sequence 1354743, A
33	49	100.0	32	31	US-09-724-059-17560	Sequence 1354743, A
34	49	100.0	32	31	US-09-724-059-17560	Sequence 1354743, A
35	49	100.0	32	31	US-09-724-059-17560	Sequence 1354743, A
36	49	100.0	32	31	US-09-724-059-17560	Sequence 1354743, A
37	49	100.0	32	31	US-09-724-059-17560	Sequence 1354743, A
38	49	100.0	32	31	US-09-724-059-17560	Sequence 1354743, A
39	49	100.0	32	31	US-09-724-059-17560	Sequence 1354743, A
40	49	100.0	32	31	US-09-724-059-17560	Sequence 1354743, A
41	49	100.0	32	31	US-09-724-059-17560	Sequence 1354743, A
42	49	100.0	32	31	US-09-724-059-17560	Sequence 1354743, A
43	49	100.0	32	31	US-09-724-059-17560	Sequence 1354743, A
44	49	100.0	32	31	US-09-724-059-17560	Sequence 1354743, A
45	49	100.0	32	31	US-09-724-059-17560	Sequence 1354743, A

ALIGNMENTS

RESULT 1
US-08-842-385-10
Sequence 10, Application US/08842385
GENERAL INFORMATION:
APPLICANT: Russell, John
TITLE OF INVENTION: Colpitts, Tracey
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
FOR DETECTING DISEASE OF THE PROSTATE
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/842,385
FILING DATE:
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Porembski, Priscilla E
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 6084.US.01
TELEPHONE: 847/937-6365
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: None
US-08-842-385-10

Query Match 100.0%; Score 49; DB 12; Length 49;
Best Local Similarity 100.0%; Pred. No. 9.9e-40;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPLGPRGDSPLLRPOHLMDOGMRHSFASGPELLRQDKRRSGSTGS 49
DB 1 EPLGPRGDSPLLRPOHLMDOGMRHSFASGPELLRQDKRRSGSTGS 49

RESULT 2
US-09-991-681-31
Sequence 31, Application US/09991681
GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA
COHEN, MAURICE
COLPITTS, TRACEY L.
FRIEDMAN, PAULA N.
GORDON, JULIAN
GRANADOS, EDWARD N.
HODGES, STEVEN C.
KLASS, MICHAEL R.
KRATOCHVIL, JON D.
ROBERTS-RAP, LISA
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSER: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/991,681
FILING DATE: 26-Nov-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/065,383
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6084.US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: None
SEQUENCE DESCRIPTION: SEQ ID NO: 31:
US-09-991-681-31

Query Match 100.0%; Score 49; DB 25; Length 49;
Best Local Similarity 100.0%; Pred. No. 9.9e-40;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPLGPRGDSPLLRPOHLMDOGMRHSFASGPELLRQDKRRSGSTGS 49
DB 1 EPLGPRGDSPLLRPOHLMDOGMRHSFASGPELLRQDKRRSGSTGS 49

RESULT 3
US-08-842-385-6
Sequence 6, Application US/08842385
GENERAL INFORMATION:
APPLICANT: Russell, John
APPLICANT: Colpitts, Tracey
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
FOR DETECTING DISEASE OF THE PROSTATE
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSER: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/842,385
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Porembski, Priscilla E
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 6084.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/937-6365
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: None
US-08-842-385-6

Query Match 100.0%; Score 49; DB 12; Length 467;
Best Local Similarity 100.0%; Pred. No. 7.4e-39;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPLGPRGDSPLLRPOHLMDOGMRHSFASGPELLRQDKRRSGSTGS 49
DB 342 EPLGPRGDSPLLRPOHLMDOGMRHSFASGPELLRQDKRRSGSTGS 390

```
RESULT 4
US-09-991-681-27
; Sequence 27, Application US/09991681
; GENERAL INFORMATION:
; APPLICANT: BILLING-MEDEL, PATRICIA
; COHEN, MAURICE
; COLPITTS, TRACEY L.
; FRIEDMAN, PAULA N.
; GORDON, JULIAN
; GRANADOS, EDWARD N.
; HODGES, STEVEN C.
; KLAAS, MICHAEL R.
; KRATOCHVIL, JON D.
; ROBERTS-RAPP, LISA
; TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
; FOR DETECTING DISEASES OF THE PROSTATE
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: PASTESEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/991,681
; FILING DATE: 26-Nov-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/065,383
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Becker, Cheryl L.
; REGISTRATION NUMBER: 35,441
; REFERENCE/DOCKET NUMBER: 6084.US.P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847/935-1729
; TELEFAX: 847/938-2623
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 518 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: None
; SEQUENCE DESCRIPTION: SEQ ID NO: 27:
US-09-991-681-27

Query Match 100.0%; Score 49; DB 25; Length 518;
Best Local Similarity 100.0%; Pred. No. 8.1e-39;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EPLGPRGDSPLLPORPOHMDQGMRHSFSAGPELLRQDKRPRSGSTGS 49
Db 393 EPLGPRGDSPLLPORPOHMDQGMRHSFSAGPELLRQDKRPRSGSTGS 441

RESULT 5
PCT-US03-01943-44
; Sequence 44, Application PC/TUS0301943
; GENERAL INFORMATION:
; APPLICANT: ORIGENE TECHNOLOGIES INC
; TITLE OF INVENTION: CANCER GENES
; FILE REFERENCE: 3U 9U 901 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/01943
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: US 10/054,935
```

```
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: US 60/356,130
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 10/102,946
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: US 10/117,229
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: US 10/144,198
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: US 10/197,824
; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44
; LENGTH: 1770
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-01943-44

Query Match 100.0%; Score 49; DB 1; Length 1770;
Best Local Similarity 100.0%; Pred. No. 2.4e-38;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EPLGPRGDSPLLPORPOHMDQGMRHSFSAGPELLRQDKRPRSGSTGS 49
Db 1645 EPLGPRGDSPLLPORPOHMDQGMRHSFSAGPELLRQDKRPRSGSTGS 1693

RESULT 6
US-10-144-198-44
; Sequence 44, Application US/10144198
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies Inc
; TITLE OF INVENTION: Regulated Prostate Cance Genes
; FILE REFERENCE: 9U 105 R1
; CURRENT APPLICATION NUMBER: US/10/144,198
; CURRENT FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 44
; LENGTH: 1770
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-198-44

Query Match 100.0%; Score 49; DB 27; Length 1770;
Best Local Similarity 100.0%; Pred. No. 2.4e-38;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EPLGPRGDSPLLPORPOHMDQGMRHSFSAGPELLRQDKRPRSGSTGS 49
Db 1645 EPLGPRGDSPLLPORPOHMDQGMRHSFSAGPELLRQDKRPRSGSTGS 1693

RESULT 7
PCT-US01-08631-40087
; Sequence 40087, Application PC/TUS0108631
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-049
; CURRENT APPLICATION NUMBER: PCT/US01/08631
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 40087
; LENGTH: 1807
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```

; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (48)..(62)
; OTHER INFORMATION: Fusion glycoprotein F0 domain identified by eMATRIX,
; OTHER INFORMATION: accession number PF00523D, p-value=7.188e-10, raw score of 11.39
; NAME/KEY: DOMAIN
; LOCATION: (941)..(950)
; OTHER INFORMATION: Helper component proteinase domain identified by Pfam,
; OTHER INFORMATION: accession name Peptidase_C6, E-value=0.0056, Pfam score of 7.9
PCT-US01-08631-40087

Query Match          100.0%; Score 49; DB 1; Length 1807;
Best Local Similarity 100.0%; Pred. No. 2.5e-38;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EPLGRGDSPLLRPOHLMDOGMHRSFSAAGPELLRQDKRRSGSTGS 49
Db 1682 EPLGRGDSPLLRPOHLMDOGMHRSFSAAGPELLRQDKRRSGSTGS 1730

RESULT 8
PCT-US01-42950-495
; Sequence 495, Application PC/TUS0142950
; GENERAL INFORMATION:
; APPLICANT: HySeq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-096
; CURRENT APPLICATION NUMBER: PCT/US01/42950
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 09/714,936
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 682
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 495
; LENGTH: 1839
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-42950-495

Query Match          100.0%; Score 49; DB 1; Length 1839;
Best Local Similarity 100.0%; Pred. No. 2.5e-38;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EPLGRGDSPLLRPOHLMDOGMHRSFSAAGPELLRQDKRRSGSTGS 49
Db 1714 EPLGRGDSPLLRPOHLMDOGMHRSFSAAGPELLRQDKRRSGSTGS 1762

RESULT 9
US-10-416-993-495
; Sequence 495, Application US/10416993
; GENERAL INFORMATION:
; APPLICANT: HySeq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-096
; CURRENT APPLICATION NUMBER: US/10/416,993
; PRIOR FILING DATE: 2003-11-16
; PRIOR APPLICATION NUMBER: 09/714,936
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 682
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 495
; LENGTH: 1839
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-416-993-495

Query Match          100.0%; Score 49; DB 30; Length 1839;
Best Local Similarity 100.0%; Pred. No. 2.5e-38;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EPLGRGDSPLLRPOHLMDOGMHRSFSAAGPELLRQDKRRSGSTGS 49
Db 1714 EPLGRGDSPLLRPOHLMDOGMHRSFSAAGPELLRQDKRRSGSTGS 1762

; DB 1714 EPLGRGDSPLLRPOHLMDOGMHRSFSAAGPELLRQDKRRSGSTGS 1762
; RESULT 10
; PCT-US03-04508-32
; Sequence 32, Application PC/TUS0304508
; GENERAL INFORMATION:
; APPLICANT: IDEC PHARMACEUTICALS
; APPLICANT: GATELY, DENNIS
; TITLE OF INVENTION: PROSTATE SPECIFIC GENES AND THE USE THEREOF IN DESIGN
; FILE REFERENCE: 037003/0301985
; CURRENT APPLICATION NUMBER: PCT/US03/04508
; CURRENT FILING DATE: 2003-02-19
; PRIOR APPLICATION NUMBER: 60/357,140
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: 60/396,082
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 60/386,759
; PRIOR FILING DATE: 2002-06-10
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 1872
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-04508-32

Query Match          100.0%; Score 49; DB 1; Length 1872;
Best Local Similarity 100.0%; Pred. No. 2.5e-38;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EPLGRGDSPLLRPOHLMDOGMHRSFSAAGPELLRQDKRRSGSTGS 49
Db 1747 EPLGRGDSPLLRPOHLMDOGMHRSFSAAGPELLRQDKRRSGSTGS 1795

RESULT 11
PCT-US01-08631-40090
; Sequence 40090, Application PC/TUS0108631
; GENERAL INFORMATION:
; APPLICANT: HySeq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-049
; CURRENT APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 40090
; LENGTH: 1982
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-08631-40090

Query Match          100.0%; Score 49; DB 1; Length 1982;
Best Local Similarity 100.0%; Pred. No. 2.7e-38;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EPLGRGDSPLLRPOHLMDOGMHRSFSAAGPELLRQDKRRSGSTGS 49
Db 1747 EPLGRGDSPLLRPOHLMDOGMHRSFSAAGPELLRQDKRRSGSTGS 1795
```

```
Db      1796 EPLGPRGDSPLORPQHLMDOGQMRHSFSAGPELLRODKRPRSGSTGS 1844

RESULT 12
PCT-US03-01943-30
; Sequence 30, Application PC/TUS0301943
; GENERAL INFORMATION:
; APPLICANT: ORIGENE TECHNOLOGIES INC
; TITLE OF INVENTION: CANCER GENES
; FILE REFERENCE: 3U 9U 901 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/01943
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: US 10/054,935
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: US 60/356,130
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 10/102,946
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: US 10/117,229
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: US 10/144,198
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: US 10/197,824
; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30
; LENGTH: 2221
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-01943-30

Query Match      100.0%; Score 49; DB 1; Length 2221;
Best Local Similarity 100.0%; Pred. No. 3e-38;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy      1 EPLGPRGDSPLORPQHLMDOGQMRHSFSAGPELLRODKRPRSGSTGS 49
Db      2096 EPLGPRGDSPLORPQHLMDOGQMRHSFSAGPELLRODKRPRSGSTGS 2144

RESULT 13
US-10-144-198-30
; Sequence 30, Application US/10144198
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies Inc
; TITLE OF INVENTION: Regulated Prostate Cance Genes
; FILE REFERENCE: 9U 105 R1
; CURRENT APPLICATION NUMBER: US/10/144,198
; CURRENT FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 30
; LENGTH: 2221
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-198-30

Query Match      100.0%; Score 49; DB 27; Length 2221;
Best Local Similarity 100.0%; Pred. No. 3e-38;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy      1 EPLGPRGDSPLORPQHLMDOGQMRHSFSAGPELLRODKRPRSGSTGS 49
Db      2096 EPLGPRGDSPLORPQHLMDOGQMRHSFSAGPELLRODKRPRSGSTGS 2144

RESULT 14
PCT-US01-00663-28663
; Sequence 28663, Application PC/TUS0100663
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharon G.

Db      1796 EPLGPRGDSPLORPQHLMDOGQMRHSFSAGPELLRODKRPRSGSTGS 1844

; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: PB 0004 WO 7
; CURRENT APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR APPLICATION NUMBER: US 09/608,408
; NUMBER OF SEQ ID NOS: 38837
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
; SEQ ID NO 28863
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007688.15
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
; OTHER INFORMATION: EST_HUMAN HIT: BE531168.1, EVALU6 3.90e-01
PCT-US01-00663-28663

Query Match      16.3%; Score 8; DB 1; Length 32;
Best Local Similarity 100.0%; Pred. No. 6.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy      27 HSFSAPE 34
Db      13 HSFSAPE 20

RESULT 15
US-09-864-761-35805
; Sequence 35805, Application US/09864761
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
```

```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 35805
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007688.15
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 35
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 31
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.9
; OTHER INFORMATION: EST_HUMAN HIT: BE531168.1, EVALUAE 3.90e-01
; US-09-864-761-35805

```

```

Query Match      16.3%; Score 8; DB 23; Length 32;
Best Local Similarity 100.0%; Pred.No. 6.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      27 HSFSAQPE 34
        |||||
        |||||
Db      13 HSFSAQPE 20

```

Search completed: July 24, 2003, 12:45:57
Job time : 74.8228 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 12:22:19 ; Search time 1.43485 Seconds
(without alignments)
140.227 Million cell updates/sec

Title: US-09-991-681-31

Perfect score: 49
Sequence: 1 EPLGPRGDSPLQRPQHLMDQGMHRSFSGAPPELLRQDKRPRSGSTGS 49

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 41799 seqs, 4106219 residues

Word size : 0

Total number of hits satisfying chosen parameters: 41799

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49	100.0	1872	US-10-367-978-32	Sequence 32, Appl
2	7	14.3	418	US-09-291-417D-99	Sequence 99, Appl
3	7	14.3	1162	US-10-294-433-224	Sequence 224, Appl
4	6	12.2	10	US-10-462-850-104	Sequence 104, Appl
5	6	12.2	166	US-10-273-573-9011	Sequence 9011, Appl
6	6	12.2	173	US-10-273-573-9548	Sequence 9548, Appl
7	6	12.2	238	US-10-273-573-6890	Sequence 6890, Appl
8	6	12.2	261	US-10-360-849A-15	Sequence 15, Appl
9	6	12.2	403	US-10-411-910A-27	Sequence 27, Appl
10	6	12.2	697	US-10-273-573-7423	Sequence 7423, Appl
11	6	12.2	729	PCT-US03-19660-37	Sequence 37, Appl
12	6	12.2	1264	US-10-273-573-7424	Sequence 7424, Appl
13	6	12.2	6	US-10-316-175-62	Sequence 62, Appl
14	5	10.2	6	US-10-316-175-63	Sequence 63, Appl
15	5	10.2	6	US-10-462-850-612	Sequence 612, Appl
16	5	10.2	10	US-10-462-850-1372	Sequence 1372, Appl
17	5	10.2	10	US-10-462-850-2092	Sequence 2092, Appl
18	5	10.2	10	US-10-462-850-2094	Sequence 2094, Appl
19	5	10.2	10	US-10-462-850-2096	Sequence 2096, Appl
20	5	10.2	10	US-10-462-850-2628	Sequence 2628, Appl
21	5	10.2	10	US-10-462-850-3103	Sequence 3103, Appl
22	5	10.2	10	US-10-462-850-3105	Sequence 3105, Appl
23	5	10.2	10	US-10-462-850-3151	Sequence 3151, Appl
24	5	10.2	60	PCT-US02-41612A-560	Sequence 560, Appl
25	5	10.2	65	US-10-273-573-9340	Sequence 9340, Appl

27	5	10.2	81	6	US-10-316-175-73	Sequence 73, Appl
28	5	10.2	86	6	US-10-273-573-7393	Sequence 7393, Appl
29	5	10.2	88	1	PCT-US02-41612A-559	Sequence 559, Appl
30	5	10.2	92	6	US-10-273-573-9360	Sequence 9360, Appl
31	5	10.2	107	1	PCT-US03-19934-119	Sequence 119, Appl
32	5	10.2	108	6	US-10-273-573-9385	Sequence 9385, Appl
33	5	10.2	110	6	US-10-273-573-9138	Sequence 9138, Appl
34	5	10.2	119	6	US-10-273-573-9512	Sequence 9512, Appl
35	5	10.2	120	6	US-10-273-573-9513	Sequence 9513, Appl
36	5	10.2	132	6	US-10-273-573-6535	Sequence 6535, Appl
37	5	10.2	135	1	PCT-US03-11231-25	Sequence 25, Appl
38	5	10.2	153	6	US-10-273-573-7166	Sequence 7166, Appl
39	5	10.2	155	6	US-10-273-573-7878	Sequence 7878, Appl
40	5	10.2	157	6	US-10-273-573-9113	Sequence 9113, Appl
41	5	10.2	161	5	US-09-635-3598-30	Sequence 30, Appl
42	5	10.2	164	6	US-10-294-433-845	Sequence 845, Appl
43	5	10.2	187	1	PCT-US02-41612A-443	Sequence 443, Appl
44	5	10.2	192	6	US-10-273-573-10544	Sequence 10544, Appl
45	5	10.2	196	6	US-10-273-573-9114	Sequence 9114, Appl

ALIGNMENTS

```
RESULT 1
US-10-367-978-32
Sequence 32, Application US/10367978
GENERAL INFORMATION:
APPLICANT: GATELEY, DENNIS
TITLE OF INVENTION: PROSTATE SPECIFIC GENES AND THE USE THEREOF IN DESIGN
FILE REFERENCE: 037003-0301988
CURRENT FILING DATE: 2003-02-19
PRIOR APPLICATION NUMBER: US/10/367,978
PRIOR FILING DATE: 2002-02-19
PRIOR APPLICATION NUMBER: 60/396,082
PRIOR FILING DATE: 2002-07-17
PRIOR APPLICATION NUMBER: 60/386,759
PRIOR FILING DATE: 2002-06-10
NUMBER OF SEQ ID NOS: 89
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 32
LENGTH: 1872
TYPE: PRT
ORGANISM: Homo sapiens
US-10-367-978-32

Query Match      100.0%; Score 49; DB 6; Length 1872;
Best Local Similarity 100.0%; Pred. No. 8.9e-39;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy      1 EPLGPRGDSPLQRPQHLMDQGMHRSFSGAPPELLRQDKRPRSGSTGS 49
Db      1747 EPLGPRGDSPLQRPQHLMDQGMHRSFSGAPPELLRQDKRPRSGSTGS 1795

RESULT 2
US-09-291-417D-99
Sequence 99, Application US/09291417D
GENERAL INFORMATION:
APPLICANT: PLOWMAN, GREGORY
APPLICANT: MARTINEZ, RICARDO
APPLICANT: WHYTE, DAVID
TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
FILE REFERENCE: 038602/0329
CURRENT FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: US/09/291,417D
PRIOR FILING DATE: 1998-04-14
NUMBER OF SEQ ID NOS: 155
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 99
```

```

; LENGTH: 418
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-291-417D-99

Query Match
Best Local Similarity 14.3%; Score 7; DB 5; Length 418;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 PELLROD 39
Db 235 PELLROD 241

RESULT 3
US-10-294-433-224
; Sequence 224, Application US/10294433
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 792CIP4
; CURRENT APPLICATION NUMBER: US/10/294,433
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: PCT/US01/14826
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/989,600
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 09/577,408
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: 10/115,831
; PRIOR FILING DATE: 2002-04-02
; PRIOR APPLICATION NUMBER: 09/677,298
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/695,781
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 10/150,802
; PRIOR FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: 09/715,869
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 10/167,379
; PRIOR FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 09/775,330
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 864
; SOFTWARE: Custom
; SEQ ID NO 224
; LENGTH: 1162
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-294-433-224

Query Match
Best Local Similarity 14.3%; Score 7; DB 6; Length 1162;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 LIRQDKR 41
Db 941 LIRQDKR 947

RESULT 4
US-10-462-850-104
; Sequence 104, Application US/10462850
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/10/462,850
; PRIOR FILING DATE: 2003-06-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProtPatent version 1.0
; SEQ ID NO 104
; LENGTH: 10

; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in BTF3L3 at 25-34 and may interact with Sequenc
US-10-462-850-104

Query Match
Best Local Similarity 12.2%; Score 6; DB 6; Length 10;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 PRSGST 47
Db 1 PRSGST 6

RESULT 5
US-10-273-573-9011
; Sequence 9011, Application US/10273573
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: US/10/273,573
; PRIOR FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 9011
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-273-573-9011

Query Match
Best Local Similarity 12.2%; Score 6; DB 6; Length 166;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 AGPELL 36
Db 109 AGPELL 114

RESULT 6
US-10-273-573-9548
; Sequence 9548, Application US/10273573
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: US/10/273,573
; PRIOR FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 9548
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(173)
; OTHER INFORMATION: xaa = X or * as defined in Table 2
US-10-273-573-9548

Query Match
Best Local Similarity 12.2%; Score 6; DB 6; Length 173;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```


Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 PLORP 16
|||||

Db 49 PLORP 54

RESULT 7
US-10-273-573-6890
; Sequence 6890, Application US/10273573
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: US/10/273,573
; CURRENT FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 6890
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(238)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-273-573-6890

Query Match 12.2%; Score 6; DB 6; Length 238;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LGPRQ 8
|||||

Db 87 LGPRQ 92

RESULT 8
US-10-360-849A-15
; Sequence 15, Application US/10360849A
; GENERAL INFORMATION:
; APPLICANT: Discovery Genomics, Inc.
; APPLICANT: Hackett, Perry
; APPLICANT: Nasevicius, Aidan
; APPLICANT: Essner, Jeffrey
; APPLICANT: Clark, Karl
; APPLICANT: Larson, Jon
; APPLICANT: Eker, Stephen
; APPLICANT: Roberg-Perez, Sharon
; TITLE OF INVENTION: FACTORS FOR ANGIOGENESIS, VASCULOGENESIS, CARTILAGE FORMATION,
; TITLE OF INVENTION: BONE FORMATION, AND METHODS OF USE THEREOF
; FILE REFERENCE: 3021.05US02
; CURRENT APPLICATION NUMBER: US/10/360,849A
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: US 60/354,978
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 15
; LENGTH: 261
; TYPE: PRT
; ORGANISM: mus musculus
US-10-360-849A-15

Query Match 12.2%; Score 6; DB 6; Length 261;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 31 AGPELL 36
|||||

Db 2 AGPELL 7

RESULT 9
US-10-360-849A-18
; Sequence 18, Application US/10360849A
; GENERAL INFORMATION:
; APPLICANT: Discovery Genomics, Inc.
; APPLICANT: Hackett, Perry
; APPLICANT: Nasevicius, Aidan
; APPLICANT: Essner, Jeffrey
; APPLICANT: Clark, Karl
; APPLICANT: Larson, Jon
; APPLICANT: Eker, Stephen
; APPLICANT: Roberg-Perez, Sharon
; TITLE OF INVENTION: FACTORS FOR ANGIOGENESIS, VASCULOGENESIS, CARTILAGE FORMATION,
; TITLE OF INVENTION: BONE FORMATION, AND METHODS OF USE THEREOF
; FILE REFERENCE: 3021.05US02
; CURRENT APPLICATION NUMBER: US/10/360,849A
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: US 60/354,978
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18
; LENGTH: 261
; TYPE: PRT
; ORGANISM: home sapiens
US-10-360-849A-18

Query Match 12.2%; Score 6; DB 6; Length 261;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 31 AGPELL 36
|||||

Db 2 AGPELL 7

RESULT 10
US-10-411-910A-27
; Sequence 27, Application US/10411910A
; GENERAL INFORMATION:
; APPLICANT: Dillon, Harrison F.
; TITLE OF INVENTION: Methods and Compositions for Evolving Hydrogenase Genes
; FILE REFERENCE: H2041203-P
; CURRENT APPLICATION NUMBER: US/10/411,910A
; CURRENT FILING DATE: 2003-04-12
; NUMBER OF SEQ ID NOS: 343
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 27
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Scenedesmus obliquus
US-10-411-910A-27

Query Match 12.2%; Score 6; DB 6; Length 403;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 42 PRSGST 47
|||||

Db 165 PRSGST 170

RESULT 11
US-10-273-573-7423
; Sequence 7423, Application US/10273573
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc

```

; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: US/10/273,573
; CURRENT FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 7423
; LENGTH: 697
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: DOMAIN
; LOCATION: (77)..(220)
; OTHER INFORMATION: Carboxyl transferase domain identified by Pfam, accession
; OTHER INFORMATION: name Carboxyl_trans, E-value=4.3e-72, Pfam score of 252.9
US-10-273-573-7423

```

```

Query Match      12.2%; Score 6; DB 6; Length 697;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      35 LIRQDK 40
      |||||
DB      442 LIRQDK 447

```

RESULT 12

PCT-US03-19660-37

```

; Sequence 37, Application PC/TUS0319660
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION; MARQUIS, Joseph P.;
; APPLICANT: BAUGHN, Mariah R.; TRAN, Uyen K.;
; APPLICANT: HAPALIA, April J.A.; KABLE, Amy E.;
; APPLICANT: EMERLING, Brooke M.; ELIHOT, Vicki S.;
; APPLICANT: LINDQUIST, Erika A.; RICHARDSON, Thomas W.;
; APPLICANT: KHARE, Reena; SWARNKAR, Anita;
; APPLICANT: LEE, Soo Yeun; RAMKUMAR, Jayalaxmi;
; APPLICANT: CHAWLA, Narinder K.; BECHA, Shanya D.;
; APPLICANT: MASON, Patricia M.; HAWKINS, Phillip R.;
; APPLICANT: BULOCH, Sean A.; JIN, Pei;
; APPLICANT: BHATIA, Umesh G.; BURRILL, John D.;
; APPLICANT: LEE, Sally; ZHENG, Wenjin;
; APPLICANT: HO, Ann; ZHENG, Wenjin;
; TITLE OF INVENTION: KINASES AND PHOSPHATASES
; FILE REFERENCE: PF-1454 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/19660
; CURRENT FILING DATE: 2003-06-19
; PRIOR APPLICATION NUMBER: US 60/390,652
; PRIOR FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/396,196
; PRIOR FILING DATE: 2002-07-15
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PERL Program
; SEQ ID NO 37
; LENGTH: 729
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 7511813CD1
PCT-US03-19660-37

```

```

Query Match      12.2%; Score 6; DB 1; Length 729;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      42 PRSGST 47
      |||||
DB      143 PRSGST 148

```

```

RESULT 13
US-10-273-573-7424
; Sequence 7424, Application US/10273573
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: US/10/273,573
; CURRENT FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 7424
; LENGTH: 1264
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-273-573-7424

```

```

Query Match      12.2%; Score 6; DB 6; Length 1264;
Best Local Similarity 100.0%; Pred. No. 11e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      35 LIRQDK 40
      |||||
DB      1020 LIRQDK 1025

```

RESULT 14

US-10-316-175-62

```

; Sequence 62, Application US/10316175
; GENERAL INFORMATION:
; APPLICANT: WELLA AG
; APPLICANT: Prass, Elon
; APPLICANT: Levy-Nissenbaum, Eitgar
; APPLICANT: Frydman, Moshe
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING HYPERTROPHIC SIMPLEX
; FILE REFERENCE: 532862000100
; CURRENT APPLICATION NUMBER: US/10/316,175
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: 60/338,188
; PRIOR FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 62
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-316-175-62

```

```

Query Match      10.2%; Score 5; DB 6; Length 6;
Best Local Similarity 100.0%; Pred. No. 3.2e+04;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      45 GSTGS 49
      |||||
DB      2 GSTGS 6

```

RESULT 15

US-10-316-175-63

```

; Sequence 63, Application US/10316175
; GENERAL INFORMATION:
; APPLICANT: WELLA AG
; APPLICANT: Prass, Elon
; APPLICANT: Levy-Nissenbaum, Eitgar
; APPLICANT: Frydman, Moshe

```

APPLICANT: Thederahn, Ted
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
TITLE OF INVENTION: AND TREATING HYPOTHYROIDISM SIMPLEX
FILE REFERENCE: 532862000100
CURRENT APPLICATION NUMBER: US/10/316,175
CURRENT FILING DATE: 2002-12-09
PRIOR APPLICATION NUMBER: 60/338,188
PRIOR FILING DATE: 2001-12-07
NUMBER OF SEQ ID NOS: 73
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 63
LENGTH: 6
TYPE: PRT
ORGANISM: Homo sapiens
US-10-316-175-63

Query Match 10.2%; Score 5; DB 6; Length 6;
Best Local Similarity 100.0%; Pred. No. 3.2e+04;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 45 GSTGS 49
DB 1 GSTGS 5

Search completed: July 24, 2003, 12:46:20
Job time : 2.63485 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OW protein - protein search, using sw model

Run on: July 24, 2003, 12:16:59 ; Search time 5.88287 Seconds
(without alignments)
801.015 Million cell updates/sec

Title: US-09-991-681-31

Perfect score: 49

Sequence: 1 EPLGPRGQDSPLQRPQHL.....SAGPELLRQDKRPRSGSTGS 49

Scoring table: OLIGO

Searched: 283308 seqs, 96168682 residues

Word size: 0

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: PIR_76:*

1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7	14.3	86	2 C95304	hypothetical trans
2	7	14.3	139	2 S23800	pathogenesis-relat
3	7	14.3	147	1 S18598	pathogenesis-relat
4	7	14.3	196	2 S41270	ribosomal protein
5	7	14.3	260	2 AG2760	conserved hypotet
6	7	14.3	269	2 E97541	hypothetical prote
7	7	14.3	334	2 S74737	conserved hypotet
8	7	14.3	415	2 C75592	conserved hypotet
9	7	14.3	433	2 S77508	extra protein - syn
10	7	14.3	439	2 AH2093	dihydroxycotase (im
11	7	14.3	1162	2 T30194	nuclear protein SA
12	7	14.3	1532	2 H96795	hypothetical prote
13	6	12.2	90	2 T24161	hypothetical prote
14	6	12.2	99	2 D97460	hypothetical prote
15	6	12.2	118	2 T15239	hypothetical prote
16	6	12.2	130	2 JQ1000	hypothetical 14.7K
17	6	12.2	130	2 C36803	hypothetical prote
18	6	12.2	138	2 T06173	photosystem II 10K
19	6	12.2	147	2 D72504	hypothetical prote
20	6	12.2	154	2 E81984	hypothetical prote
21	6	12.2	154	2 A81039	conserved hypotet
22	6	12.2	163	2 A48770	Ca2+-transporting
23	6	12.2	175	2 RKWT59	ribulose-bisphosph
24	6	12.2	176	2 S61845	probable transcrip
25	6	12.2	192	2 A81753	Orf50 (bacterioph
26	6	12.2	194	2 A87333	hypothetical prote
27	6	12.2	195	2 D81030	transcription regu
28	6	12.2	205	2 D95200	conserved hypotet
29	6	12.2	205	2 C98067	hypothetical prote

30	6	12.2	210	2 B44984	collagen - nematod
31	6	12.2	215	1 SXADMS	hexon-associated p
32	6	12.2	233	2 AC2587	conserved hypotet
33	6	12.2	233	2 C97369	hypothetical prote
34	6	12.2	234	2 D83403	hypothetical prote
35	6	12.2	234	2 A82814	chitin-phosphate
36	6	12.2	237	2 S64057	probable membrane
37	6	12.2	266	2 T09872	endo-1,3-beta-gluc
38	6	12.2	274	2 C75255	iron ABC transport
39	6	12.2	276	2 T12477	hypothetical prote
40	6	12.2	283	2 T32921	hypothetical prote
41	6	12.2	290	2 S66072	yabg protein - Bac
42	6	12.2	297	2 D87011	conserved hypotet
43	6	12.2	297	2 T18638	hypothetical prote
44	6	12.2	298	2 S53483	probable membrane
45	6	12.2	300	2 T42343	hypothetical prote

ALIGNMENTS

RESULT 1
C95304
hypothetical transposase, partial match [imported] - Sinorhizobium meliloti (strain 1021)
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C:Accession: C95304
R:Barnett, M.J., Fisher, R.F., Jones, T., Komp, C., Abola, A.P., Barloy-Hubler, F., Bows,
; Kalman, S., Keating, D.H., Palm, C., Peck, M.C., Surzycki, R., Weiss, D.H., Yeh, K.C.
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilot
A:Reference number: A95262; MUID:21396509; PMID:11481432
A:Accession: C95304
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-86 <KUR>
A:Cross-references: GB:AB006469; PIDN:AAK64997.1; PID:G14523425; GSPDB:GN00165
A:Experimental source: strain 1021, megaplasmid pSymA
R:Galibert, F., Finan, T.M., Long, S.R., Puhler, A., Abola, P., Ampe, F., Barloy-Hubler,
pela, D., Chain, P., Cowie, A., Davis, R.W., Dreano, S., Federspiel, N.A., Fisher, R.F.,
L., Hyman, R.W., Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D., Kahn, M.L., Kalman, S., Keating, D.H., Kiss, E., Komp, C., Lelaure,
hehault, P., Vandenbol, M., Voeholter, F.J., Weidner, S., Wells, D.H., Wong, K., Yeh, K.
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
C:Genetics:
A:Contents: annotation
A:Gene: Sma0643
A:Genome: plasmid

Query Match 14.3%; Score 7; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Or 31 AGPELLR 37
Db 66 AGPELLR 72

RESULT 2
S23800
pathogenesis-related protein 4B - common tobacco
C:Species: Nicotiana tabacum (common tobacco)
C:Date: 22-Nov-1993 #sequence_revision 01-Sep-1995 #text_change 23-Mar-2001
C:Accession: S23800
R:Linthorst, H.J.M., Danhaess, N., Brederode, F.T., van Kan, J.A.L., de Wit, P.J.G.M., Bo
Mol. Plant Microbe Interact. 4, 586-592, 1991
A:Title: Tobacco and tomato PR proteins homologous to win and pro-hevein lack the "hevein
A:Reference number: S23799; MUID:2208317; PMID:1804403
A:Accession: S23800
A:Status: preliminary
A:Molecule type: mRNA

A:Residues: 1-139 <LIN>
 A:Cross-references: EMBL:X58547; NID:g19965; PIDN:CAA4138.1; PID:g19966
 C:Superfamily: pathogenesis-related protein 4A; barwin homology
 F:18-139/Domain: barwin homology <BAR>

Query Match
 Best Local Similarity 14.3%; Score 7; DB 2; Length 139;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GPRGDS 10
 DB 71 GPRGDS 77

RESULT 3

S18598
 pathogenesis-related protein 4 - common tobacco
 C:Species: Nicotiana tabacum (common tobacco)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: S18598
 R:Friedrich, L.; Moyer, M.; Ward, E.; Ryals, J.
 Mol. Gen. Genet. 230, 113-119, 1991
 A:Title: Pathogenesis-related protein 4 is structurally homologous to the carboxy-termini
 A:Reference number: S18598; MUID:92079884; PMID:1745223
 A:Accession: S18598
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-147 <FRI>
 A:Cross-references: GB:X60281; NID:g19963; PIDN:CAA42820.1; PID:g19964
 C:Superfamily: pathogenesis-related protein 4A; barwin homology
 F:26-147/Domain: barwin homology <BAR>

Query Match
 Best Local Similarity 14.3%; Score 7; DB 1; Length 147;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GPRGDS 10
 DB 79 GPRGDS 85

RESULT 4

S41270
 ribosomal protein S4, chloroplast - Iris pallida chloroplast (fragment)
 C:Species: chloroplast Iris pallida
 C:Date: 19-Mar-1997 #sequence_revision 06-Jun-1997 #text_change 13-Aug-1999
 C:Accession: S41270
 R:Nadot, S.; Bajon, R.; Lejeune, B.
 submitted to the EMBL Data Library, January 1994
 A:Description: Chloroplast rps4 gene as a tool for the study of Poaceae phylogeny.
 A:Reference number: S41254
 A:Accession: S41270
 A:Molecule type: DNA
 A:Residues: 1-196 <NAD>
 A:Cross-references: EMBL:Z29254; NID:g439945; PIDN:CAA82453.1; PID:g439946
 C:Genetics:
 A:Gene: rps4
 A:Genome: chloroplast
 C:Superfamily: Escherichia coli ribosomal protein S4
 C:Keywords: chloroplast; protein biosynthesis; ribosome

Query Match
 Best Local Similarity 14.3%; Score 7; DB 2; Length 196;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 KRPRGS 46
 DB 24 KRPRGS 30

RESULT 5

AG2760
 conserved hypothetical protein Atu1495 [imported] - Agrobacterium tumefaciens (strain C5

C:Species: Agrobacterium tumefaciens

C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
 C:Accession: AG2760
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
 ; Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Kreppan, W.; Perry, M.; Gordon-Kamm,
 ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; MUID:21608550; PMID:11743193

A:Accession: AG2760

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-260 <KUR>

A:Cross-references: GB:AE006688; PIDN:AA142501.1; PID:g17739919; GSPDB:GN00186

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: Atu1495

A:Map position: circular chromosome

C:Superfamily: hypothetical protein HI0454

Query Match
 Best Local Similarity 14.3%; Score 7; DB 2; Length 260;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 FSAGPEL 35
 DB 157 FSAGPEL 163

RESULT 6

E97541
 hypothetical protein AGR_C 2756 [imported] - Agrobacterium tumefaciens (strain C58, Cere
 C:Species: Agrobacterium tumefaciens
 C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
 C:Accession: E97541
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
 Science 294, 2323-2328, 2001
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
 A:Reference number: A97359; MUID:21608551; PMID:11743194
 A:Accession: E97541
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-269 <KUR>
 A:Cross-references: GB:AE007869; PIDN:AAK87286.1; PID:g15156578; GSPDB:GN00169
 C:Genetics:
 A:Gene: AGR_C 2756
 A:Map position: circular chromosome
 C:Superfamily: hypothetical protein HI0454

Query Match
 Best Local Similarity 14.3%; Score 7; DB 2; Length 269;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 FSAGPEL 35
 DB 166 FSAGPEL 172

RESULT 7

S74737
 conserved hypothetical protein s110926 - Synecchocystis sp. (strain PCC 6803)
 C:Species: Synecchocystis sp.
 A:Variety: PCC 6803
 C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
 C:Accession: S74737
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
 O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
 DNA Res. 3, 109-136, 1996
 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecchocystis
 sp.

A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S74737
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-334 <KAN>
A:Cross-references: EMBL:D90901; GB:AB001339; NID:g1651897; PIDN:BAI16888.1; PID:g165196
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
A:Superfamily: Haemophilus influenzae conserved hypothetical protein HI0634

Query Match 14.3%; Score 7; DB 2; Length 334;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 30 SAGPELL 36
Db 313 SAGPELL 319

RESULT 8

C75592
conserved hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: C75592
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; Ma
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: C75592
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-415 <WMI>
A:Cross-references: GB:AB001863; GB:AB001825; NID:g6460670; PIDN:AAI12424.1; PID:g646071
A:Experimental source: strain R1
A:Genetics:
A:Gene: DRA0368
A:Map position: 2

Query Match 14.3%; Score 7; DB 2; Length 415;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 QRPCHLM 20
Db 49 QRPCHLM 55

RESULT 9

S77508
atrra protein - Synechocystis sp. (strain PCC 6803)
N:Alternate names: protein B11897
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
C:Accession: S77508
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O. K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-116, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
sp.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S77508
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-433 <KAN>
A:Cross-references: EMBL:D90905; GB:AB001339; NID:g1652360; PIDN:BAI17355.1; PID:d101808
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
A:Genetics:
A:Gene: atrra

Query Match 14.3%; Score 7; DB 2; Length 433;

Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 30 SAGPELL 36
Db 273 SAGPELL 279

RESULT 10

AH2093
dihydroorotase [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C:Accession: AH2093
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriyuchi
Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AH2093
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-439 <KUR>
A:Cross-references: GB:BA000019; PIDN:BAE74002.1; PID:g17131395; GSPDB:GN00179
A:Experimental source: strain PCC 7120
A:Genetics:
A:Gene: all2303
A:Superfamily: Bacillus dihydroorotase; Bacillus dihydroorotase homology

Query Match 14.3%; Score 7; DB 2; Length 439;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 34 ELFRQDK 40
Db 242 ELFRQDK 248

RESULT 11

T30194
nuclear protein SA2 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 20-Jun-2000
C:Accession: T30194
R:Barbero, J.L.; Carreiro, C.
submitted to the EMBL Data Library, November 1997
A:Description: Homologue to human nuclear protein SA2.
A:Reference number: Z20769
A:Accession: T30194
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1162 <BAR>
A:Cross-references: EMBL:AJ002636; PIDN:CA05638.1
A:Experimental source: lung
A:Genetics:
A:Gene: SAP2

Query Match 14.3%; Score 7; DB 2; Length 1162;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 35 LLRQDKR 41
Db 941 LLRQDKR 947

RESULT 12

H96795
hypothetical protein F28016.11 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: H96795

R.Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Hansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luo, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, Ker, M.; Wu, D.; Yu, G.; Frazer, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: H96795
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1532 <STO>
 A:Cross-References: GB:AE005173; NID:g6143891; PIDN:AAF04437.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: F28016.11
 A:Map position: 1

Query Match 14.3%; Score 7; DB 2; Length 1532;
 Best Local Similarity 100.0%; Pred. No. 60;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 RSGSTGS 49
 |||||
 Db 146 RSGSTGS 152

RESULT 13
 T24161
 hypothetical protein R11.2 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 16-Feb-2000
 C:Accession: T24161
 A:Reference number: Z19847
 A:Accession: T24161
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-90 <MTL>
 A:Cross-References: EMBL:Z81577; PIDN:CAB04650.1; GSPDB:GN00028; CESP:R11.2
 A:Experimental source: clone R11
 C:Genetics:
 A:Gene: CESP:R11.2
 A:Map position: X
 A:introns: 33/2

Query Match 12.2%; Score 6; DB 2; Length 90;
 Best Local Similarity 100.0%; Pred. No. 48;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 LLROD 39
 |||||
 Db 53 LLROD 58

RESULT 14
 D97460
 hypothetical protein AGR C 1518 [imported] - Agrobacterium tumefaciens (strain C58, Cere
 C:Species: Agrobacterium tumefaciens
 C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
 C:Accession: D97460
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;
 Science 294, 2323-2328, 2001
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
 A:Reference number: A97359; MUID:21608551; PMID:11743194
 A:Accession: D97460
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-99 <KUR>

A:Cross-References: GB:AE007869; PIDN:AAK6637.1; PID:g15155815; GSPDB:GN00169
 C:Genetics:
 A:Gene: AGR C 1518
 A:Map position: circular chromosome

Query Match 12.2%; Score 6; DB 2; Length 99;
 Best Local Similarity 100.0%; Pred. No. 52;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps -0;

QY 35 LLRODK 40
 |||||
 Db 22 LLRODK 27

RESULT 15
 T15239
 hypothetical protein F56F4.7 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
 C:Accession: T15239
 R:Murray, J.; Wohldmann, P.; Gillam, B.
 A:Submitted to the EMBL Data Library, May 1997
 A:Description: The sequence of C. elegans cosmid F56F4.
 A:Reference number: Z18313
 A:Accession: T15239
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-118 <MDR>
 A:Cross-References: EMBL:AF003146; NID:g2088769; PID:g2088766; PIDN:AAB54204.1; GSPDB:GN
 A:Experimental source: strain Bristol N2; clone F56F4
 C:Genetics:
 A:Gene: CESP:F56F4.7
 A:Map position: 1
 A:introns: 58/1; 82/2

Query Match 12.2%; Score 6; DB 2; Length 118;
 Best Local Similarity 100.0%; Pred. No. 61;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 SGSTGS 49
 |||||
 Db 59 SGSTGS 64

Search completed: July 24, 2003, 12:27:22
 Job time : 6.88287 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 12:15:23 ; Search time 3.2284 Seconds

(without alignments)
713.761 Million cell updates/sec

Title: US-09-991-681-31

Sequence: 1 EPLGPGQDSPLQRPQHLM.....SAGPELLRQDRPRSGSTGS 49

Scoring table: OLIGO

Gapop 60.0 , Gapept 60.0

Searched: 127863 seqs, 47026705 residues

Word size : 0

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Swissprot_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7	14.3	147	1	PR4A_TOBAC
2	7	14.3	147	1	PR4B_TOBAC
3	7	14.3	182	1	RR4_BELCH
4	7	14.3	182	1	RR4_HEAMA
5	7	14.3	182	1	RR4_HYMLI
6	7	14.3	182	1	RR4_IRITU
7	7	14.3	182	1	RR4_ISOTA
8	7	14.3	182	1	RR4_LIBFO
9	7	14.3	182	1	RR4_NEOSP
10	7	14.3	183	1	RR4_NERBO
11	7	14.3	183	1	RR4_ARICA
12	7	14.3	183	1	RR4_BABST
13	7	14.3	183	1	RR4_CRONU
14	7	14.3	183	1	RR4_CROSP
15	7	14.3	183	1	RR4_FRESE
16	7	14.3	183	1	RR4_LAPNE
17	7	14.3	183	1	RR4_SPASP
18	7	14.3	194	1	RR4_BOBGL
19	7	14.3	194	1	RR4_CYPSP
20	7	14.3	194	1	RR4_FURGI
21	7	14.3	194	1	RR4_IREN
22	7	14.3	194	1	RR4_MORST
23	7	14.3	194	1	RR4_SISST
24	7	14.3	194	1	RR4_TRISA
25	7	14.3	195	1	RR4_GLAMP
26	7	14.3	195	1	RR4_PILTE
27	7	14.3	195	1	RR4_ROMRE
28	7	14.3	195	1	RR4_WATAN
29	7	14.3	196	1	RR4_AGABR
30	7	14.3	196	1	RR4_ASASB
31	7	14.3	196	1	RR4_TRIPA
32	7	14.3	196	1	RR4_NAROD
33	7	14.3	197	1	RR4_GLAPA

34	7	14.3	214	1	BT33_HUMAN	Q13892 homo sapien
35	7	14.3	934	1	V926_SYNY3	P72872 synechocyst
36	7	14.3	953	1	B3A4_RAT	O8K4V2 rattus norv
37	7	14.3	955	1	B3A4_RABIT	O9GKY1 oryctolagus
38	7	14.3	983	1	B3A4_HUMAN	O9GQ91 homo sapien
39	7	14.3	1162	1	SA2_HUMAN	O8N3U4 homo sapien
40	7	14.3	1162	1	SA2_MOUSE	O35638 mus muscullu
41	6	12.2	125	1	VNIU_AZOB	O43895 azospirillu
42	6	12.2	130	1	VG75_HAVER	P28985 equine herp
43	6	12.2	138	1	PSBR_HORVU	O40070 hordeum vul
44	6	12.2	154	1	Y645_NEIMA	O9JW77 neisseria m
45	6	12.2	154	1	Y116_NEIMB	O9JXZ9 neisseria m

ALIGNMENTS

```

RESULT 1
PR4A_TOBAC
ID PR4A_TOBAC STANDARD; PRT; 147 AA.
AC P29062;
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Pachogenesis-related protein PR-4A precursor.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; Lamiales; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Xanthi; TISSUE=leaf;
RX MEDLINE=92079884; PubMed=174523;
RA Friedrich L., Moyer W., Ward B., Ryals J.;
RT "Pachogenesis-related protein 4 is structurally homologous to the
RT carboxy-terminal domains of hevein, win-1 and win-2.";
RL Mol. Gen. Genet. 230:113-119(1991).
CC -!- INDUCTION: By TMV infection.
CC -!- SIMILARITY: BELONGS TO THE BARWIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X60281; CAA42820.1; -.
CC PIR; S18598; S18598.
CC HSSP; P28814; 1BM4.
CC InterPro; IPR001153; Barwin.
CC Pfam; PF00967; Barwin; 1.
CC PRINTS; PR00602; BARWIN.
CC PRODOM; PD004535; Barwin; 1.
CC PROSITE; PS00771; BARWIN_1; 1.
CC PROSITE; PS00772; BARWIN_2; 1.
CC Plant defense; Pachogenesis-related protein; Signal.
CC SIGNAL
FT CHAIN 1 25
FT DISULFID 54 86
FT DISULFID 75 109
FT DISULFID 89 145
SQ SEQUENCE 147 AA; 16221 MW; 800602A2AFB40D30 CRC64;

Query Match 14.3%; Score 7; DB 1; Length 147;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY 4 GPRGDS 10
DB 79 GPRGDS 85

```

```

RESULT 2
PRAB_TOBAC STANDARD; PRT; 147 AA.
ID PRAB_TOBAC
AC P29063;
RT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Patchogenesis-related protein PR-4B precursor.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamiales; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Xanthi; TISSUE=Leaf;
RX MEDLINE=92079884; PubMed=1745223;
RA Friedrich L., Moyer M., Ward E., Ryals J.;
RT "pathogenesis-related protein 4 is structurally homologous to the
RT carboxy-terminal domains of hevein, Wtn-1 and Wtn-2.";
RL Mol. Gen. Genet. 230:113-119(1991).
CC -1- INDUCTION: BY TMV infection.
CC -1- SIMILARITY: BELONGS TO THE BARWIN FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X60282; CAA42821.1; -
DR HSSP; P28814; 1BW4.
DR InterPro; IPR001153; Barwin.
DR Pfam; PF00967; Barwin.1.
DR PRINTS; PR00602; BARWIN.1.
DR PRODOM; PD004535; Barwin.1.
DR PROSITE; PS00771; BARWIN_1; 1.
DR PROSITE; PS00772; BARWIN_2; 1.
KW Plant defense; Patchogenesis-related protein; Signal.
FT SIGNAL 1 25
FT CHAIN 26 147 PATHOGENESIS-RELATED PROTEIN PR-4B.
FT DISULFID 54 86 BY SIMILARITY.
FT DISULFID 75 109 BY SIMILARITY.
FT DISULFID 89 145 BY SIMILARITY.
SQ SEQUENCE 147 AA; 16235 MW; FEF7579E1FB2C874 CRC64;

Query Match 14.3%; Score 7; DB 1; Length 147;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GPRGDS 10
DB 79 GPRGDS 85

RESULT 3
RR4_BELCH STANDARD; PRT; 182 AA.
ID RR4_BELCH
AC O19590;
RT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chloroplast 30S ribosomal protein S4 (Fragment).
GN RPS4.
OS Belamcanda chinensis.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Iridaceae;
OC Belamcanda.

```

```

OX NCBI_TaxID=58944;
RN [1]
RP SEQUENCE FROM N.A.
RA Souza-Chies T.T., Bittar G., Nadot S., Carter L., Besin E.,
RA Lejeune B.P.;
RT "Phylogenetic analysis of Iridaceae with parsimony and distance
RT methods using the plastid gene rps4.";
RL Plant Syst. Evol. 204:109-123(1997).
CC -1- FUNCTION: One of the primary rRNA binding proteins, it binds
CC directly to 16S rRNA where it nucleates assembly of the body of
CC the 30S subunit (By similarity).
CC -1- FUNCTION: With S5 and S12 plays an important role in translational
CC accuracy (By similarity).
CC -1- SUBUNIT: Part of the 30S ribosomal subunit. Contacts protein S5.
CC The interaction surface between S4 and S5 is involved in control
CC of translational fidelity (By similarity).
CC -1- SUBCELLULAR LOCATION: Chloroplast.
CC -1- SIMILARITY: Contains 1 S4 RNA-binding domain.
CC -1- SIMILARITY: BELONGS TO THE S4P FAMILY OF RIBOSOMAL PROTEINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; Z68235; CAA92533.1; -
DR HSSP; P81288; 1C05.
DR HAMAP; MF_01306; -; 1.
DR InterPro; IPR001912; Ribosomal_S4.
DR InterPro; IPR002942; S4.
DR InterPro; IPR005709; S4_bact_ORG.
DR Pfam; PF00163; Ribosomal_S4; 1.
DR Pfam; PF01479; S4; 1.
DR SMART; SM00363; S4; 1.
DR TIGRPFAMs; TIGR01017; rpsd_bact; 1.
DR PROSITE; PS00632; RIBOSOMAL_S4; 1.
DR PROSITE; PS50889; S4; 1.
KW Ribosomal protein; RNA-binding; rRNA-binding; Chloroplast.
FT NON TER 1 1
FT DOMAIN 82 143 S4 RNA-BINDING.
FT NON TER 182 182
SQ SEQUENCE 182 AA; 21048 MW; E3620B0641A04EFS CRC64;

Query Match 14.3%; Score 7; DB 1; Length 182;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 KRPRGS 46
DB 17 KRPRGS 23

RESULT 4
RR4_HEAMA STANDARD; PRT; 182 AA.
ID RR4_HEAMA
AC O20231;
RT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chloroplast 30S ribosomal protein S4 (Fragment).
GN RPS4.
OS Haemanthus magnificus.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Amaryllidaceae;
OC Haemanthus.
OX NCBI_TaxID=59038;
RN [1]
RP SEQUENCE FROM N.A.
RA Souza-Chies T.T., Bittar G., Nadot S., Carter L., Besin E.,

```

```

RA Lejeune B.P.;
RT "Phylogenetic analysis of Iridaceae with parsimony and distance
RL Plant Syst. Evol. 204:109-123(1997).
CC -1- FUNCTION: One of the primary RNA binding proteins, it binds
CC directly to 16S rRNA where it nucleates assembly of the body of
CC the 30S subunit (By similarity).
CC -1- FUNCTION: With S5 and S12 plays an important role in translational
CC accuracy (By similarity).
CC -1- SUBUNIT: Part of the 30S ribosomal subunit. Contacts protein S5.
CC The interaction surface between S4 and S5 is involved in control
CC of translational fidelity (By similarity).
CC -1- SUBCELLULAR LOCATION: Chloroplast.
CC -1- SIMILARITY: Contains 1 S4 RNA-binding domain.
CC -1- SIMILARITY: BELONGS TO THE S4P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z68253; CAA92551.1; -.
CC HAMAP; MF_01306; -.
CC HSSP; P81288; IC05.
CC InterPro; IPR001912; Ribosomal_S4.
CC InterPro; IPR002942; S4.
CC InterPro; IPR005709; S4_bact_org.
CC Pfam; PF00163; Ribosomal_S4; 1.
CC Pfam; PF01479; S4; 1.
CC SMART; SM00363; S4; 1.
CC TIGRfam; TIGR01017; rpsd_bact; 1.
CC PROSITE; PS00632; RIBOSOMAL_S4; 1.
CC PROSITE; PS50889; S4; 1.
CC Ribosomal protein; RNA-binding; RNA-binding; Chloroplast.
FT DOMAIN 1 82 143 S4 RNA-BINDING.
FT NON_TER 182 182
FT NON_TER 182 182
SQ SEQUENCE 182 AA; 21086 MW; E091810DPFD6D3B CRC64;

Query Match 14.3%; Score 7; DB 1; Length 182;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 KRPRSGS 46
DB 17 KRPRSGS 23

RESULT 5
ID RR4_HYMLI STANDARD; PRT; 182 AA.
AC O20230;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chloroplast 30S ribosomal protein S4 (Fragment).
GN RPS4.
OS Hymenocallis littoralis (Beach spiderlily).
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Amaryllidaceae;
OC Hydnocallis.
OX NCBI_TaxID=59040;
RN SEQUENCE FROM N.A.
RA Souza-Chies T.T., Bittar G., Nadot S., Carter L., Besin E.,
RT Lejeune B.P.;
RL "Phylogenetic analysis of Iridaceae with parsimony and distance
methods using the plastid gene rps4."
PL Plant Syst. Evol. 204:109-123(1997).

```

```

CC -1- FUNCTION: One of the primary RNA binding proteins, it binds
CC directly to 16S rRNA where it nucleates assembly of the body of
CC the 30S subunit (By similarity).
CC -1- FUNCTION: With S5 and S12 plays an important role in translational
CC accuracy (By similarity).
CC -1- SUBUNIT: Part of the 30S ribosomal subunit. Contacts protein S5.
CC The interaction surface between S4 and S5 is involved in control
CC of translational fidelity (By similarity).
CC -1- SUBCELLULAR LOCATION: Chloroplast.
CC -1- SIMILARITY: Contains 1 S4 RNA-binding domain.
CC -1- SIMILARITY: BELONGS TO THE S4P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z68251; CAA92549.1; -.
CC HAMAP; MF_01306; -.
CC HSSP; P81288; IC05.
CC InterPro; IPR001912; Ribosomal_S4.
CC InterPro; IPR002942; S4.
CC InterPro; IPR005709; S4_bact_org.
CC Pfam; PF00163; Ribosomal_S4; 1.
CC Pfam; PF01479; S4; 1.
CC SMART; SM00363; S4; 1.
CC TIGRfam; TIGR01017; rpsd_bact; 1.
CC PROSITE; PS00632; RIBOSOMAL_S4; 1.
CC PROSITE; PS50889; S4; 1.
CC Ribosomal protein; RNA-binding; RNA-binding; Chloroplast.
FT DOMAIN 1 82 143 S4 RNA-BINDING.
FT NON_TER 182 182
FT NON_TER 182 182
SQ SEQUENCE 182 AA; 21050 MW; 78B1261C3F6DC8AF CRC64;

Query Match 14.3%; Score 7; DB 1; Length 182;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 KRPRSGS 46
DB 17 KRPRSGS 23

RESULT 6
ID RR4_IRILU STANDARD; PRT; 182 AA.
AC O20235;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chloroplast 30S ribosomal protein S4 (Fragment).
GN RPS4.
OS Iris lutescens.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Iridaceae;
OC Iris.
OX NCBI_TaxID=58956;
RN SEQUENCE FROM N.A.
RA Souza-Chies T.T., Bittar G., Nadot S., Carter L., Besin E.,
RT Lejeune B.P.;
RL "Phylogenetic analysis of Iridaceae with parsimony and distance
methods using the plastid gene rps4."
PL Plant Syst. Evol. 204:109-123(1997).
CC -1- FUNCTION: One of the primary RNA binding proteins, it binds
CC directly to 16S rRNA where it nucleates assembly of the body of
CC the 30S subunit (By similarity).
CC -1- FUNCTION: With S5 and S12 plays an important role in translational

```

```

CC accuracy (By similarity).
CC - SUBUNIT: Part of the 30S ribosomal subunit. Contacts protein S5.
CC The interaction surface between S4 and S5 is involved in control
CC of translational fidelity (By similarity).
CC - SUBCELLULAR LOCATION: Chloroplast.
CC - SIMILARITY: Contains 1 S4 RNA-binding domain.
CC - SIMILARITY: BELONGS TO THE S4P FAMILY OF RIBOSOMAL PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z68242; CAA92540.1; -.
CC HSSP: P81288; IC05.
CC HAMAP: MF_01306; -.
CC InterPro: IPR001912; Ribosomal_S4.
CC InterPro: IPR002942; S4.
CC InterPro: IPR005709; S4_bact_org.
CC Pfam: PF00163; Ribosomal_S4; 1.
CC Pfam: PF01479; S4; 1.
CC SMART: SM00363; S4; 1.
CC SMART: TIGR01017; rpsd_bact; 1.
CC TIGRPFAM: TIGR01017; rpsd_bact; 1.
CC PROSITE: PS00632; RIBOSOMAL_S4; 1.
CC PROSITE: PS50889; S4; 1.
CC Ribosomal protein; RNA-binding; Chloroplast.
CC NON TER 1 1
CC DOMAIN 82 143 S4 RNA-BINDING.
CC FT NON TER 182 182
CC FT DOMAIN 82 143
CC FT NON TER 182 182
CC SQ SEQUENCE 182 AA; 20996 MW; 193292203A86C6C6 CRC64;

Query Match
Best Local Similarity 14.3%; Score 7; DB 1; Length 182;
Matches 7; Conservative 100.0%; Pred. No. 4.3;
Mismatches 0; Indels 0; Gaps 0;

QY 40 KRPRS GS 46
DB 17 KRPRS GS 23

RESULT 7
RR4_ISOTA STANDARD; PRT; 182 AA.
ID RR4_ISOTA
AC 020336;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chloroplast 30S ribosomal protein S4 (Fragment).
GN RPS4.
OS Isophysis tasmanica.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Iridaceae;
OC Isophysis.
OC NCB1_TaxID=49752;
RN NCB1_TaxID=49752;
RP SEQUENCE FROM N.A.
RA Souza-Chies T.T., Bittar G., Nadot S., Carter L., Besin E.,
RA Lejeune B.P.;
RT "Phylogenetic analysis of Iridaceae with parsimony and distance
RT methods using the plastid gene rps4."
RL Plant Syst. Evol. 204:109-123(1997).
CC - FUNCTION: One of the primary rRNA binding proteins, it binds
CC directly to 16S rRNA where it nucleates assembly of the body of
CC the 30S subunit (By similarity).
CC - FUNCTION: With S5 and S12 plays an important role in translational
CC accuracy (By similarity).
CC - SUBUNIT: Part of the 30S ribosomal subunit. Contacts protein S5.
CC The interaction surface between S4 and S5 is involved in control
CC of translational fidelity (By similarity).

```

```

CC - SUBCELLULAR LOCATION: Chloroplast.
CC - SIMILARITY: Contains 1 S4 RNA-binding domain.
CC - SIMILARITY: BELONGS TO THE S4P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z68243; CAA92541.1; -.
CC HSSP: P81288; IC05.
CC HAMAP: MF_01306; -.
CC InterPro: IPR001912; Ribosomal_S4.
CC InterPro: IPR002942; S4.
CC InterPro: IPR005709; S4_bact_org.
CC Pfam: PF00163; Ribosomal_S4; 1.
CC Pfam: PF01479; S4; 1.
CC SMART: SM00363; S4; 1.
CC SMART: TIGR01017; rpsd_bact; 1.
CC PROSITE: PS00632; RIBOSOMAL_S4; 1.
CC PROSITE: PS50889; S4; 1.
CC Ribosomal protein; RNA-binding; rRNA-binding; Chloroplast.
CC NON TER 1 1
CC DOMAIN 82 143 S4 RNA-BINDING.
CC FT NON TER 182 182
CC FT DOMAIN 82 143
CC FT NON TER 182 182
CC SQ SEQUENCE 182 AA; 21040 MW; D716DC01669E0213 CRC64;

Query Match
Best Local Similarity 14.3%; Score 7; DB 1; Length 182;
Matches 7; Conservative 100.0%; Pred. No. 4.3;
Mismatches 0; Indels 0; Gaps 0;

QY 40 KRPRS GS 46
DB 17 KRPRS GS 23

RESULT 8
RR4_LIBFO STANDARD; PRT; 182 AA.
ID RR4_LIBFO
AC 020254;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chloroplast 30S ribosomal protein S4 (Fragment).
GN RPS4.
OS Libertia formosa (Snowy merald).
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Iridaceae;
OC Libertia.
OC NCB1_TaxID=58960;
RN NCB1_TaxID=58960;
RP SEQUENCE FROM N.A.
RA Souza-Chies T.T., Bittar G., Nadot S., Carter L., Besin E.,
RA Lejeune B.P.;
RT "Phylogenetic analysis of Iridaceae with parsimony and distance
RT methods using the plastid gene rps4."
RL Plant Syst. Evol. 204:109-123(1997).
CC - FUNCTION: One of the primary rRNA binding proteins, it binds
CC directly to 16S rRNA where it nucleates assembly of the body of
CC the 30S subunit (By similarity).
CC - FUNCTION: With S5 and S12 plays an important role in translational
CC accuracy (By similarity).
CC - SUBUNIT: Part of the 30S ribosomal subunit. Contacts protein S5.
CC The interaction surface between S4 and S5 is involved in control
CC of translational fidelity (By similarity).
CC - SUBCELLULAR LOCATION: Chloroplast.
CC - SIMILARITY: Contains 1 S4 RNA-binding domain.
CC - SIMILARITY: BELONGS TO THE S4P FAMILY OF RIBOSOMAL PROTEINS.

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; Z68245; CAA92543.1; -
DR HSSP; P81288; IC05.
DR HAMAP; MF_01306; -; 1.
DR InterPro; IPR001912; Ribosomal_S4.
DR InterPro; IPR002942; S4.
DR InterPro; IPR005709; S4_bact.org.
DR Pfam; PF00163; Ribosomal_S4; 1.
DR SMART; SM00363; S4; 1.
DR SMART; SM00363; S4; 1.
DR TIGRfam; TIGR01017; rpsd_bact; 1.
DR PROSITE; PS00632; RIBOSOMAL_S4; 1.
DR PROSITE; PS50889; S4; 1.
DR Ribosomal protein; RNA-binding; rRNA-binding; Chloroplast.
FT NON TER 1 1
FT DOMAIN 82 143 S4 RNA-BINDING.
FT NON TER 182 182
SQ SEQUENCE 182 AA; 20963 MW; 6AECA864142FC0 CRC64;

Query Match 14.3%; Score 7; DB 1; Length 182;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 40 KRPRSGS 46
Db 17 KRPRSGS 23

RESULT 9
RR4_NEOSP STANDARD; PRT; 182 AA.
ID RR4_NEOSP
AC 036051;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chloroplast 30S ribosomal protein S4 (Fragment).
GN RPS4.
OS Neomartia sp. (strain Lejeune 1997).
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Iridaceae;
OC Neomartia.
OX NCBI_TaxID=58966;
RN [1]
RP SEQUENCE FROM N.A.
RA Souza-Chies T.T., Bittar G., Nadot S., Carter L., Besin E.,
RA Lejeune B.P.;
RT "Phylogenetic analysis of Iridaceae with parsimony and distance
RT methods using the plastid gene rps4.";
RL Plant Syst. Evol. 204:109-123(1997).
CC -1- FUNCTION: One of the primary rRNA binding proteins, it binds
CC directly to 16S rRNA where it nucleates assembly of the body of
CC the 30S subunit (By similarity).
CC -1- FUNCTION: With S5 and S12 plays an important role in translational
CC accuracy (By similarity).
CC -1- SUBUNIT: Part of the 30S ribosomal subunit. Contacts protein S5.
CC The interaction surface between S4 and S5 is involved in control
CC of translational fidelity (By similarity).
CC -1- SUBCELLULAR LOCATION: Chloroplast.
CC -1- SIMILARITY: BELONGS TO THE S4P FAMILY OF RIBOSOMAL PROTEINS.
CC -1- SIMILARITY: BELONGS TO THE S4P FAMILY OF RIBOSOMAL PROTEINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; Z68247; CAA92545.1; -
DR HSSP; P81288; IC05.
DR HAMAP; MF_01306; -; 1.
DR InterPro; IPR001912; Ribosomal_S4.
DR InterPro; IPR002942; S4.
DR InterPro; IPR005709; S4_bact.org.
DR Pfam; PF00163; Ribosomal_S4; 1.
DR SMART; SM00363; S4; 1.
DR SMART; SM00363; S4; 1.
DR TIGRfam; TIGR01017; rpsd_bact; 1.
DR PROSITE; PS00632; RIBOSOMAL_S4; 1.
DR PROSITE; PS50889; S4; 1.
DR Ribosomal protein; RNA-binding; rRNA-binding; Chloroplast.
FT NON TER 1 1
FT DOMAIN 82 143 S4 RNA-BINDING.
FT NON TER 182 182
SQ SEQUENCE 182 AA; 20993 MW; B6FDA4E2D06BC903 CRC64;

Query Match 14.3%; Score 7; DB 1; Length 182;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 40 KRPRSGS 46
Db 17 KRPRSGS 23

RESULT 10
RR4_NERBO STANDARD; PRT; 182 AA.
ID RR4_NERBO
AC 020265;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chloroplast 30S ribosomal protein S4 (Fragment).
GN RPS4.
OS Nerine bowdenii (Guernsey 111y).
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Amaryllidaceae;
OC Nerine.
OX NCBI_TaxID=59042;
RN [1]
RP SEQUENCE FROM N.A.
RA Souza-Chies T.T., Bittar G., Nadot S., Carter L., Besin E.,
RA Lejeune B.P.;
RT "Phylogenetic analysis of Iridaceae with parsimony and distance
RT methods using the plastid gene rps4.";
RL Plant Syst. Evol. 204:109-123(1997).
CC -1- FUNCTION: One of the primary rRNA binding proteins, it binds
CC directly to 16S rRNA where it nucleates assembly of the body of
CC the 30S subunit (By similarity).
CC -1- FUNCTION: With S5 and S12 plays an important role in translational
CC accuracy (By similarity).
CC -1- SUBUNIT: Part of the 30S ribosomal subunit. Contacts protein S5.
CC The interaction surface between S4 and S5 is involved in control
CC of translational fidelity (By similarity).
CC -1- SUBCELLULAR LOCATION: Chloroplast.
CC -1- SIMILARITY: BELONGS TO THE S4P FAMILY OF RIBOSOMAL PROTEINS.
CC -1- SIMILARITY: BELONGS TO THE S4P FAMILY OF RIBOSOMAL PROTEINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: Z68252; CAA92550.1; -.
 DR HSSP; P81288; IC05.
 DR HAMAP; MF_01306; -; 1.
 DR InterPro; IPR001912; Ribosomal_S4.
 DR InterPro; IPR002942; S4.
 DR InterPro; IPR005709; S4_bact_org.
 DR Pfam; PF00163; Ribosomal_S4; 1.
 DR Pfam; PF01479; S4; 1.
 DR SMART; SM00363; S4; 1.
 DR TIGRfams; TIGR01017; rpsd_bact; 1.
 DR PROSITE; PS00632; RIBOSOMAL_S4; 1.
 DR PROSITE; PS50889; S4; 1.
 DR Ribosomal protein; RNA-binding; Chloroplast.
 FT NON TER 1 1
 FT DOMAIN 82 143 S4 RNA-BINDING.
 FT NON TER 182 182
 SQ SEQUENCE 182 AA; 21054 MW; 96CEDF24A44E5CCD CRC64;

Query Match 14.3%; Score 7; DB 1; Length 182;
 Best Local Similarity 100.0%; Pred. No. 4.3;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 KRPRS GS 46
 |||||
 Db 17 KRPRS GS 23

RESULT 11
 RR4_ARICA STANDARD; PRT; 183 AA.
 ID RR4_ARICA 019983;
 AC 019815; 019983; (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Chloroplast 30S ribosomal protein S4 (Fragment).
 GN RPS4.
 OS Aristeae capitata, and
 OS Aristeae platycaulis.
 OG Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Iridaceae;
 OC Aristeae.
 OC NCBI_TaxID=58983; 58984;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Souza-Chies T.T., Bittar G., Nadot S., Carter L., Besin E.,
 RA Lejeune B.P.;
 RT "Phylogenetic analysis of Iridaceae with parsimony and distance
 RT methods using the plastid gene rps4.";
 RL Plant Syst. Evol. 204:109-123(1997).
 CC -!- FUNCTION: One of the primary RNA binding proteins, it binds
 CC directly to 16S rRNA where it nucleates assembly of the body of
 CC the 30S subunit (By similarity).
 CC -!- FUNCTION: With S5 and S12 plays an important role in translational
 CC accuracy (By similarity).
 CC -!- SUBUNIT: Part of the 30S ribosomal subunit. Contacts protein S5.
 CC The interaction surface between S4 and S5 is involved in control
 CC of translational fidelity (By similarity).
 CC -!- SUBCELLULAR LOCATION: Chloroplast.
 CC -!- SIMILARITY: Contains 1 S4 RNA-binding domain.
 CC -!- SIMILARITY: BELONGS TO THE S4P FAMILY OF RIBOSOMAL PROTEINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Z68233; CAA92531.1; -.
 CC EMBL; Z68232; CAA92530.1; -.
 CC HSSP; P81288; IC05.

DR HAMAP; MF_01306; -; 1.
 DR InterPro; IPR001912; Ribosomal_S4.
 DR InterPro; IPR002942; S4.
 DR InterPro; IPR005709; S4_bact_org.
 DR Pfam; PF00163; Ribosomal_S4; 1.
 DR Pfam; PF01479; S4; 1.
 DR SMART; SM00363; S4; 1.
 DR TIGRfams; TIGR01017; rpsd_bact; 1.
 DR PROSITE; PS00632; RIBOSOMAL_S4; 1.
 DR PROSITE; PS50889; S4; 1.
 DR Ribosomal protein; RNA-binding; Chloroplast.
 FT NON TER 1 1
 FT DOMAIN 82 143 S4 RNA-BINDING.
 FT NON TER 183 183
 SQ SEQUENCE 183 AA; 21254 MW; 3686C0140C384CT1 CRC64;

Query Match 14.3%; Score 7; DB 1; Length 183;
 Best Local Similarity 100.0%; Pred. No. 4.3;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 KRPRS GS 46
 |||||
 Db 17 KRPRS GS 23

RESULT 12
 RR4_BABST STANDARD; PRT; 183 AA.
 ID RR4_BABST 019995;
 AC 019995;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Chloroplast 30S ribosomal protein S4 (Fragment).
 GN RPS4.
 OS Babiana sericea (Baboon flower).
 OG Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Iridaceae;
 OC Babiana.
 OC NCBI_TaxID=58942;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Souza-Chies T.T., Bittar G., Nadot S., Carter L., Besin E.,
 RA Lejeune B.P.;
 RT "Phylogenetic analysis of Iridaceae with parsimony and distance
 RT methods using the plastid gene rps4.";
 RL Plant Syst. Evol. 204:109-123(1997).
 CC -!- FUNCTION: One of the primary RNA binding proteins, it binds
 CC directly to 16S rRNA where it nucleates assembly of the body of
 CC the 30S subunit (By similarity).
 CC -!- FUNCTION: With S5 and S12 plays an important role in translational
 CC accuracy (By similarity).
 CC -!- SUBUNIT: Part of the 30S ribosomal subunit. Contacts protein S5.
 CC The interaction surface between S4 and S5 is involved in control
 CC of translational fidelity (By similarity).
 CC -!- SUBCELLULAR LOCATION: Chloroplast.
 CC -!- SIMILARITY: Contains 1 S4 RNA-binding domain.
 CC -!- SIMILARITY: BELONGS TO THE S4P FAMILY OF RIBOSOMAL PROTEINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Z68234; CAA92532.1; -.
 CC HSSP; P81288; IC05.
 DR HAMAP; MF_01306; -; 1.
 DR InterPro; IPR001912; Ribosomal_S4.
 DR InterPro; IPR002942; S4.
 DR InterPro; IPR005709; S4_bact_org.

DR Pfam; PF00163; Ribosomal_S4; 1.
 DR Pfam; PF01479; S4; 1.
 DR SMART; SM00363; S4; 1.
 DR TIGRFAMs; TIGR01017; rpsd_bact; 1.
 DR PROSITE; PS00632; RIBOSOMAL_S4; 1.
 DR PROSITE; PS50889; S4; 1.
 DR KMW Ribosomal protein; RNA-binding; Chloroplast.
 FT NON TER 1 1
 FT DOMAIN 82 143 S4 RNA-BINDING.
 FT NON TER 183 183
 SQ SEQUENCE 183 AA; 21138 MW; A7AD0354908BEAD3 CRC64;

Query Match 14.3%; Score 7; DB 1; Length 183;
 Best Local Similarity 100.0%; Pred. No. 4.3;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 40 KRPRSGS 46
 Db 17 KRPRSGS 23

RESULT 13

RR4_CRONU STANDARD; PRT; 183 AA.
 AC 020100; 020216; 020261;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Chloroplast 30S ribosomal protein S4 (Fragment).
 GN RPS4.
 OS Crocus nudiflorus,
 OS Gladiolus communis, and
 OS Micranthus juncea.
 OG Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Iridaceae;
 OC Crocus.
 OX NCBI_TaxID=58950, 58992, 58962;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Souza-Chies T.T., Bittar G., Nadot S., Carter L., Besin E.,
 RA Lejeune B.P.;
 RT "Phylogenetic analysis of Iridaceae with parsimony and distance
 RT methods using the plastid gene rps4.";
 RL Plant Syst. Evol. 204:109-123(1997).
 CC -1- FUNCTION: One of the primary rRNA binding proteins, it binds
 CC directly to 16S rRNA where it nucleates assembly of the body of
 CC the 30S subunit (By similarity).
 CC -1- FUNCTION: With S5 and S12 plays an important role in translational
 CC accuracy (By similarity).
 CC -1- SUBUNIT: Part of the 30S ribosomal subunit. Contacts protein S5.
 CC The interaction surface between S4 and S5 is involved in control
 CC of translational fidelity (By similarity).
 CC -1- SUBCELLULAR LOCATION: Chloroplast.
 CC -1- SIMILARITY: BELONGS TO THE S4P FAMILY OF RIBOSOMAL PROTEINS.
 CC -1- SIMILARITY: BELONGS TO THE S4P FAMILY OF RIBOSOMAL PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@ib-sib.ch).

DR EMBL; Z68237; CA92535.1; -;
 DR EMBL; Z68241; CA92539.1; -;
 DR EMBL; Z68246; CA92544.1; -;
 DR HSSP; P81288; IC05;
 DR HAMAP; MF_01306; -; 1.
 DR InterPro; IPR001912; Ribosomal_S4.
 DR InterPro; IPR002942; S4.
 DR InterPro; IPR005709; S4_bact_org.
 DR InterPro; IPR005709; S4_bact_org.

DR Pfam; PF00163; Ribosomal_S4; 1.
 DR Pfam; PF01479; S4; 1.
 DR SMART; SM00363; S4; 1.
 DR TIGRFAMs; TIGR01017; rpsd_bact; 1.
 DR PROSITE; PS00632; RIBOSOMAL_S4; 1.
 DR PROSITE; PS50889; S4; 1.
 DR KMW Ribosomal protein; RNA-binding; Chloroplast.
 FT NON TER 1 1
 FT DOMAIN 82 143 S4 RNA-BINDING.
 FT NON TER 183 183
 SQ SEQUENCE 183 AA; 21138 MW; A7AD03548C79EAD3 CRC64;

Query Match 14.3%; Score 7; DB 1; Length 183;
 Best Local Similarity 100.0%; Pred. No. 4.3;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 40 KRPRSGS 46
 Db 17 KRPRSGS 23

RESULT 14

RR4_CROSP STANDARD; PRT; 183 AA.
 AC 020103;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Chloroplast 30S ribosomal protein S4 (Fragment).
 GN RPS4.
 OS Crocosmia sp. (strain Lejeune 1997).
 OG Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Iridaceae;
 OC Crocosmia.
 OX NCBI_TaxID=58948;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Souza-Chies T.T., Bittar G., Nadot S., Carter L., Besin E.,
 RA Lejeune B.P.;
 RT "Phylogenetic analysis of Iridaceae with parsimony and distance
 RT methods using the plastid gene rps4.";
 RL Plant Syst. Evol. 204:109-123(1997).
 CC -1- FUNCTION: One of the primary rRNA binding proteins, it binds
 CC directly to 16S rRNA where it nucleates assembly of the body of
 CC the 30S subunit (By similarity).
 CC -1- FUNCTION: With S5 and S12 plays an important role in translational
 CC accuracy (By similarity).
 CC -1- SUBUNIT: Part of the 30S ribosomal subunit. Contacts protein S5.
 CC The interaction surface between S4 and S5 is involved in control
 CC of translational fidelity (By similarity).
 CC -1- SUBCELLULAR LOCATION: Chloroplast.
 CC -1- SIMILARITY: BELONGS TO THE S4P FAMILY OF RIBOSOMAL PROTEINS.
 CC -1- SIMILARITY: BELONGS TO THE S4P FAMILY OF RIBOSOMAL PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@ib-sib.ch).

DR EMBL; Z68236; CA92534.1; -;
 DR HSSP; P81288; IC05;
 DR HAMAP; MF_01306; -; 1.
 DR InterPro; IPR001912; Ribosomal_S4.
 DR InterPro; IPR002942; S4.
 DR InterPro; IPR005709; S4_bact_org.
 DR Pfam; PF00163; Ribosomal_S4; 1.
 DR Pfam; PF01479; S4; 1.
 DR SMART; SM00363; S4; 1.
 DR TIGRFAMs; TIGR01017; rpsd_bact; 1.

DR PROSITE: PS00632; RIBOSOMAL_S4; 1.
 DR PROSITE: PS50889; S4; 1.
 KW Ribosomal protein; RNA-binding; rRNA-binding; Chloroplast.
 FT NON_TER 1 1
 FT DOMAIN 82 143 S4 RNA-BINDING.
 FT NON_TER 183 183
 SQ SEQUENCE 183 AA; 21124 MW; A7ADDD35251CF3763 CRC64;

Query Match 14.3%; Score 7; DB 1; Length 183;
 Best Local Similarity 100.0%; Pred. No. 4.3;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 40 KRPRSGS 46
 DB 17 KRPRSGS 23

RESULT 15

RR4_PRESF STANDARD; PRT; 183 AA.
 ID RR4_PRESF
 AC 020215;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Chloroplast 30S ribosomal protein S4 (Fragment).
 GN RPS4.
 OS Friesia sp. (strain Lejeune 1997).
 OG Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Iridaceae;
 OC Friesia.
 OX NCBI_TaxID=58989;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Souza-Chies T.T., Bittar G., Nadot S., Carter L., Besin E.,
 RA Lejeune B.P.;
 RT "Phylogenetic analysis of Iridaceae with parsimony and distance
 RT methods using the plastid gene rps4.";
 RL Plant Syst. Evol. 204:109-123(1997).
 CC -1- FUNCTION: One of the primary rRNA binding proteins, it binds
 CC directly to 16S rRNA where it nucleates assembly of the body of
 CC the 30S subunit (By similarity).
 CC -1- FUNCTION: With S5 and S12 plays an important role in translational
 CC accuracy (By similarity).
 CC -1- SUBUNIT: Part of the 30S ribosomal subunit. Contacts protein S5.
 CC The interaction surface between S4 and S5 is involved in control
 CC of translational fidelity (By similarity).
 CC -1- SUBCELLULAR LOCATION: Chloroplast.
 CC -1- SIMILARITY: Contains 1 S4 RNA-binding domain.
 CC -1- SIMILARITY: BELONGS TO THE S4P FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Z68240; CAA92538.1; -
 CC HSSP: P81288; IC05.
 DR HAMAP: MF_01306; -; 1.
 DR InterPro: IPR001912; Ribosomal_S4.
 DR InterPro: IPR002942; S4.
 DR InterPro: IPR005709; S4_bact_org.
 DR Pfam: PF00163; Ribosomal_S4; 1.
 DR Pfam: PF01479; S4; 1.
 DR SMART: SM00363; S4; 1.
 DR TIGRFAMs: TIGR01017; rpsd_bact; 1.
 DR PROSITE: PS00632; RIBOSOMAL_S4; 1.
 DR PROSITE: PS50889; S4; 1.
 KW Ribosomal protein; RNA-binding; rRNA-binding; Chloroplast.
 FT NON_TER 1 1

FT DOMAIN 82 143 S4 RNA-BINDING.
 FT NON_TER 183 183
 SQ SEQUENCE 183 AA; 21143 MW; A7B96C19B1CF3763 CRC64;
 Query Match 14.3%; Score 7; DB 1; Length 183;
 Best Local Similarity 100.0%; Pred. No. 4.3;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 KRPRSGS 46
 DB 17 KRPRSGS 23
 Search completed: July 24, 2003, 12:22:11
 Job time : 4.2284 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 12:16:28 ; Search time 15.0659 Seconds
(without alignment)
839.285 Million cell updates/sec

Title: US-09-991-681-31

Perfect score: 49
Sequence: 1 EPLGPRGQDSPLQRPQHLMDGOMRHSFSAGPELLRQDKRPRSGSTGS 49

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size: 0

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_protein:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriophage:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	49	100.0	589	4	Q8N4Y4	Q8N4Y4 homo sapien
2	49	100.0	592	4	Q96CH9	Q96CH9 homo sapien
3	49	100.0	1770	4	Q9ULH6	Q9ULH6 homo sapien
4	14.3	86	16	Q922V9	Q922V9 rhizobium	
5	7	14.3	174	8	Q8WKP7	Q8WKP7 zephyra ele
6	7	14.3	174	8	Q8WKQ0	Q8WKQ0 walleria ma
7	7	14.3	174	8	Q8WKR5	Q8WKR5 recophila
8	7	14.3	175	8	Q8WLA8	Q8WLA8 doryanthes
9	7	14.3	175	8	Q8WKS8	Q8WKS8 syringodea
10	7	14.3	176	8	Q8WKN8	Q8WKN8 chasmanthe
11	7	14.3	177	8	Q8WLN8	Q8WLN8 hesperantha
12	7	14.3	182	8	Q8WKS0	Q8WKS0 solenomeius
13	7	14.3	182	8	Q8WLN4	Q8WLN4 gymandritis
14	7	14.3	182	8	Q8WLN4	Q8WLN4 ferraria cr
15	7	14.3	182	8	Q8WLN4	Q8WLN4 diplarthena
16	7	14.3	182	8	Q8WKN3	Q8WKN3 cipura camp

17	7	14.3	182	8 Q8WKV7	Q8WKV7 orthrosanth
18	7	14.3	182	8 Q8WLN3	Q8WLN3 galaxia sp.
19	7	14.3	182	8 Q8WKV6	Q8WKV6 olivium fi
20	7	14.3	183	8 Q8WLN20	Q8WLN20 ixia laetifo
21	7	14.3	183	8 Q8WKU9	Q8WKU9 pardanthops
22	7	14.3	183	8 Q8WKQ3	Q8WKQ3 tritoniopsis
23	7	14.3	183	8 Q8WKQ6	Q8WKQ6 thecianthu
24	7	14.3	183	8 Q8WKQ1	Q8WKQ1 watsonia an
25	7	14.3	190	8 Q95C87	Q95C87 chloranthus
26	7	14.3	190	8 Q95C88	Q95C88 austroball
27	7	14.3	196	8 Q9TJW7	Q9TJW7 dioscorea b
28	7	14.3	196	8 Q9TJX8	Q9TJX8 aechea dea
29	7	14.3	196	8 Q9TJW4	Q9TJW4 electaria c
30	7	14.3	196	8 Q9TJX2	Q9TJX2 costus luca
31	7	14.3	196	8 Q9TJX9	Q9TJX9 typha angus
32	7	14.3	196	8 Q9TJX3	Q9TJX3 butomus umb
33	7	14.3	196	8 Q9TJW1	Q9TJW1 musa x para
34	7	14.3	196	8 Q9TJW0	Q9TJW0 sagittaria
35	7	14.3	196	8 Q9TJX7	Q9TJX7 arum italic
36	7	14.3	196	8 Q9TJX6	Q9TJX6 allisma plan
37	7	14.3	196	8 Q9TJW3	Q9TJW3 laurus nobi
38	7	14.3	196	8 Q9TJW8	Q9TJW8 dioscorea b
39	7	14.3	202	11 Q8BXT0	Q8BXT0 mus musculu
40	7	14.3	211	2 Q9AN07	Q9AN07 bradyrhizob
41	7	14.3	269	16 Q8UPA3	Q8UPA3 agrobacteri
42	7	14.3	279	2 Q8VNL9	Q8VNL9 enterococcu
43	7	14.3	280	4 Q9PIL0	Q9PIL0 homo sapien
44	7	14.3	394	16 Q8PCT6	Q8PCT6 xanthomonas
45	7	14.3	394	16 Q8PIL0	Q8PIL0 xanthomonas

ALIGNMENTS

RESULT 1

ID Q8N4Y4 PRELIMINARY; PRT; 589 AA.

AC Q8N4Y4;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Similar to KIAA1244 protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RX [1]
RN SEQUENCE FROM N.A.
RP TISSUE=Colon;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC033191; AAH33191.1; -.
FT NON_TER
SQ SEQUENCE 589 AA; 66086 MW; 8041BEA348DE65F7 CRC64;

Query Match 100.0%; Score 49; DB 4; Length 589;
Best Local Similarity 100.0%; Pred. No. 2e-42;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EPLGPRGQDSPLQRPQHLMDGOMRHSFSAGPELLRQDKRPRSGSTGS 49
Db 464 EPLGPRGQDSPLQRPQHLMDGOMRHSFSAGPELLRQDKRPRSGSTGS 512

RESULT 2

ID Q96CH9 PRELIMINARY; PRT; 592 AA.

AC Q96CH9;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strusberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC014227; AAH14227.1; -.
KW Hypothetical protein.
FT NON TER
SQ SEQUENCE 592 AA; 66400 MW; F4A1E807B0DF47B5 CRC64;

Query Match
Best Local Similarity 100.0%; Score 49; DB 4; Length 592;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EPLGRRGDSPLLRPOHLMDOGMHRSAGPELLRQDKRRSGSGTGS 49
Db 467 EPLGRRGDSPLLRPOHLMDOGMHRSAGPELLRQDKRRSGSGTGS 515

RESULT 3
O9UH6 PRELIMINARY; PRT; 1770 AA.
AC O9UH6; Q96P46;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE BIG3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RA Hong W.;
RT "K1A1244 as a novel distantly related member (BIG3) of the BIG/Sec7p
RL subfamily of ARF GEFs."
DR EMBL; AF413080; AAL04174.1; -.
DR InterPro; IPR000904; Sec7.
DR SMART; SM00222; Sec7; 1.
KW Hypothetical protein.
SQ SEQUENCE 1770 AA; 195845 MW; 5E96E36A6F92AB4 CRC64;

Query Match
Best Local Similarity 100.0%; Score 49; DB 4; Length 1770;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EPLGRRGDSPLLRPOHLMDOGMHRSAGPELLRQDKRRSGSGTGS 49
Db 1645 EPLGRRGDSPLLRPOHLMDOGMHRSAGPELLRQDKRRSGSGTGS 1693

RESULT 4
O92ZV9 PRELIMINARY; PRT; 86 AA.
AC O92ZV9;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Hypothetical transposase, partial match.
GN RA0339 OR SMA0643.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396509; PubMed=11481432;

```

```

RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,
RA Barloy-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,
RA Gurjal M., Hong A., Huzar L., Hyman R.W., Kahn D., Kahn M.L.,
RA Kaiman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
RA Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;
RT "Nucleotide sequence and predicted functions of the entire
RT Sinorhizobium meliloti pSymA megaplasmid."
RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).
DR EMBL; AE007226; AAK64997.1; -.
KW Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 86 AA; 9653 MW; 30C5C96876F8F98 CRC64;

Query Match
Best Local Similarity 14.3%; Score 7; DB 16; Length 86;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 31 AGPELLR 37
Db 66 AGPELLR 72

RESULT 5
O8WK7 PRELIMINARY; PRT; 174 AA.
AC O8WK7;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Ribosomal protein S4 (Fragment).
GN RPS4.
OS Zephyra elegans.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales;
OC Tecophilaeaceae; Zephyra.
OX NCBI_TaxID=85344;
RN (1)
RP SEQUENCE FROM N.A.
RA Reeves G., Chase M.W., Goldblatt P., Rudall P., Fay M.F., Cox A.V.,
RA Lejeune B., Souza-Chies T.;
RT "Molecular systematics of Iridaceae: Evidence from four plastid DNA
RT regions."
RL Am. J. Bot. 88:2074-2087(2001).
DR EMBL; AJ409052; CAC83931.1; -.
DR InterPro; IPR001812; Ribosomal_S4.
DR InterPro; IPR002942; S4.
DR Pfam; PF00163; Ribosomal_S4; 1.
DR Pfam; PF01479; S4; 1.
DR SMART; SM00363; S4; 1.
DR TIGRFAMs; TIGR01017; rpsd_bact; 1.
DR PROSITE; PS00632; RIBOSOMAL_S4; 1.
KW Chloroplast.
FT NON TER
SQ SEQUENCE 174 AA; 19912 MW; 96BF6433C8CF3B5 CRC64;

Query Match
Best Local Similarity 10.0%; Score 7; DB 8; Length 174;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 40 KRPRSGS 46
Db 9 KRPRSGS 15

RESULT 6
O8WK0 PRELIMINARY; PRT; 174 AA.
AC O8WK0;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)

```

DE Ribosomal protein S4 (Fragment).
 GN RPS4.
 OS Walleria mackenzii.
 OG Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales;
 OC Tecophilaeaceae; Walleria.
 OX NCBI_TaxID=85342;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Reeves G., Chase M.W., Goldblatt P., Rudall P., Fay M.F., Cox A.V.,
 RA Lejeune B., Souza-Chies T.;
 RT "Molecular systematics of Iridaceae: Evidence from four plastid DNA
 RT regions.";
 RL Am. J. Bot. 88:2074-2087(2001).
 DR EMBL; AJ409048; CAC83927.1; -;
 DR InterPro; IPR001912; Ribosomal_S4.
 DR InterPro; IPR002942; S4.
 DR Pfam; PF00163; Ribosomal_S4; 1.
 DR Pfam; PF01479; S4; 1.
 DR SMART; SM00363; S4; 1.
 DR TIGRFAMs; TIGR01017; rpsd_bact; 1.
 DR PROSITE; PS00632; RIBOSOMAL_S4; 1.
 KW Chloroplast.
 FT NON_TER 1
 FT NON_TER 174
 SQ SEQUENCE 174 AA; 19907 MW; 0C505CB389B1F27F CRC64;

Query Match 14.3%; Score 7; DB 8; Length 174;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 40 KRPRSGS 46
 |||||
 Db 9 KRPRSGS 15

RESULT 7

OS O8MKRS PRELIMINARY; PRT; 174 AA.
 AC O8MKRS;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Ribosomal protein S4 (Fragment).
 GN RPS4.
 OS Tecophilaeae cyanococcus.
 OG Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales;
 OC Tecophilaeaceae; Tecophilaeae.
 OX NCBI_TaxID=49737;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Reeves G., Chase M.W., Goldblatt P., Rudall P., Fay M.F., Cox A.V.,
 RA Lejeune B., Souza-Chies T.;
 RT "Molecular systematics of Iridaceae: Evidence from four plastid DNA
 RT regions.";
 RL Am. J. Bot. 88:2074-2087(2001).
 DR EMBL; AJ409049; CAC83928.1; -;
 DR InterPro; IPR001912; Ribosomal_S4.
 DR InterPro; IPR002942; S4.
 DR InterPro; IPR005709; S4_bact_org.
 DR Pfam; PF00163; Ribosomal_S4; 1.
 DR Pfam; PF01479; S4; 1.
 DR TIGRFAMs; TIGR01017; rpsd_bact; 1.
 DR PROSITE; PS00632; RIBOSOMAL_S4; 1.
 KW Chloroplast.
 FT NON_TER 1
 FT NON_TER 174
 SQ SEQUENCE 174 AA; 19906 MW; 7A8B2A7389B1F26F CRC64;

Query Match 14.3%; Score 7; DB 8; Length 174;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 40 KRPRSGS 46
 |||||
 Db 9 KRPRSGS 15

RESULT 8

OS O8WL48 PRELIMINARY; PRT; 174 AA.
 AC O8WL48;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Ribosomal protein S4 (Fragment).
 GN RPS4.
 OS Doryanthaceae excelisa.
 OG Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Doryanthaceae;
 OC Doryanthaceae.
 OX NCBI_TaxID=49713;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Reeves G., Chase M.W., Goldblatt P., Rudall P., Fay M.F., Cox A.V.,
 RA Lejeune B., Souza-Chies T.;
 RT "Molecular systematics of Iridaceae: Evidence from four plastid DNA
 RT regions.";
 RL Am. J. Bot. 88:2074-2087(2001).
 DR EMBL; AJ409053; CAC83932.1; -;
 DR InterPro; IPR001912; Ribosomal_S4.
 DR InterPro; IPR002942; S4.
 DR InterPro; IPR005709; S4_bact_org.
 DR Pfam; PF00163; Ribosomal_S4; 1.
 DR Pfam; PF01479; S4; 1.
 DR SMART; SM00363; S4; 1.
 DR TIGRFAMs; TIGR01017; rpsd_bact; 1.
 DR PROSITE; PS00632; RIBOSOMAL_S4; 1.
 KW Chloroplast.
 FT NON_TER 1
 FT NON_TER 174
 SQ SEQUENCE 174 AA; 19906 MW; 7A8B2A7389B1F26F CRC64;

Query Match 14.3%; Score 7; DB 8; Length 174;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 40 KRPRSGS 46
 |||||
 Db 9 KRPRSGS 15

RESULT 9

OS O8WKS8 PRELIMINARY; PRT; 175 AA.
 AC O8WKS8;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Ribosomal protein S4 (Fragment).
 GN RPS4.
 OS Syringodea bifucata.
 OG Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Iridaceae;
 OC Syringodeae.
 OX NCBI_TaxID=152751;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Reeves G., Chase M.W., Goldblatt P., Rudall P., Fay M.F., Cox A.V.,
 RA Lejeune B., Souza-Chies T.;

```

RT "Molecular systematics of Iridaceae: Evidence from four plastid DNA
RT regions."
RL Am. J. Bot. 88:2074-2087(2001).
DR EMBL; AJ409026; CAC83844.1; -.
DR InterPro; IPR001912; Ribosomal_S4.
DR InterPro; IPR002942; S4.
DR InterPro; IPR005709; S4_bact_org.
DR Pfam; PF00163; Ribosomal_S4; 1.
DR Pfam; PF01479; S4; 1.
DR SMART; SM00363; S4; 1.
DR TIGRFAMs; TIGR01017; rpsd_bact; 1.
DR PROSITE; PS00632; RIBOSOMAL_S4; 1.
KW Chloroplast.
FT NON_TER 1
FT NON_TER 175
SQ SEQUENCE 175 AA; 19997 MW; 7DB37813CDC54854 CRC64;

Query Match 14.3%; Score 7; DB 8; Length 175;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 KRPRS GS 46
DB 9 KRPRS GS 15

RESULT 10
Q8WKX8 PRELIMINARY; PRT; 176 AA.
ID Q8WKX8;
AC Q8WKX8;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Ribosomal protein S4 (Fragment).
GN RPS4.
OS Chaetanthe aethiopica.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Iridaceae;
OC Chaetanthe.
OX NCBI_TaxID=152713;
RN [1]
RP SEQUENCE FROM N.A.
RA Reeves G., Chase M.W., Goldblatt P., Rudall P., Fay M.F., Cox A.V.,
RA Lejeune B., Souza-Chies T.;
RT "Molecular systematics of Iridaceae: Evidence from four plastid DNA
regions."
RL Am. J. Bot. 88:2074-2087(2001).
DR EMBL; AJ409022; CAC83795.1; -.
DR InterPro; IPR001912; Ribosomal_S4.
DR InterPro; IPR002942; S4.
DR InterPro; IPR005709; S4_bact_org.
DR Pfam; PF00163; Ribosomal_S4; 1.
DR Pfam; PF01479; S4; 1.
DR SMART; SM00363; S4; 1.
DR TIGRFAMs; TIGR01017; rpsd_bact; 1.
DR PROSITE; PS00632; RIBOSOMAL_S4; 1.
KW Chloroplast.
FT NON_TER 1
FT NON_TER 176
SQ SEQUENCE 176 AA; 20118 MW; 9DAC3CBDB5FE090E CRC64;

Query Match 14.3%; Score 7; DB 8; Length 176;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 KRPRS GS 46
DB 10 KRPRS GS 16

RESULT 11
Q8WML25

```

```

ID Q8WML25 PRELIMINARY; PRT; 177 AA.
AC Q8WML25;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Ribosomal protein S4 (Fragment).
GN RPS4.
OS Hesperantha pseudopilosa.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Iridaceae;
OC Hesperantha.
OX NCBI_TaxID=152728;
RN [1]
RP SEQUENCE FROM N.A.
RA Reeves G., Chase M.W., Goldblatt P., Rudall P., Fay M.F., Cox A.V.,
RA Lejeune B., Souza-Chies T.;
RT "Molecular systematics of Iridaceae: Evidence from four plastid DNA
regions."
RL Am. J. Bot. 88:2074-2087(2001).
DR EMBL; AJ409019; CAC79546.1; -.
DR InterPro; IPR001912; Ribosomal_S4.
DR InterPro; IPR002942; S4.
DR Pfam; PF00163; Ribosomal_S4; 1.
DR Pfam; PF01479; S4; 1.
DR SMART; SM00363; S4; 1.
DR PROSITE; PS00632; RIBOSOMAL_S4; 1.
KW Chloroplast.
FT NON_TER 1
FT NON_TER 177
SQ SEQUENCE 177 AA; 20169 MW; 0D857B2887F3B694 CRC64;

Query Match 14.3%; Score 7; DB 8; Length 177;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 KRPRS GS 46
DB 11 KRPRS GS 17

RESULT 12
Q8WKS0 PRELIMINARY; PRT; 182 AA.
ID Q8WKS0;
AC Q8WKS0;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Ribosomal protein S4 (Fragment).
GN RPS4.
OS Solanum elaeagnifolium.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Iridaceae;
OC Solanum.
OX NCBI_TaxID=152748;
RN [1]
RP SEQUENCE FROM N.A.
RA Reeves G., Chase M.W., Goldblatt P., Rudall P., Fay M.F., Cox A.V.,
RA Lejeune B., Souza-Chies T.;
RT "Molecular systematics of Iridaceae: Evidence from four plastid DNA
regions."
RL Am. J. Bot. 88:2074-2087(2001).
DR EMBL; AJ409038; CAC83920.1; -.
DR InterPro; IPR001912; Ribosomal_S4.
DR InterPro; IPR002942; S4.
DR InterPro; IPR005709; S4_bact_org.
DR Pfam; PF00163; Ribosomal_S4; 1.
DR Pfam; PF01479; S4; 1.
DR SMART; SM00363; S4; 1.
DR TIGRFAMs; TIGR01017; rpsd_bact; 1.
DR PROSITE; PS00632; RIBOSOMAL_S4; 1.
KW Chloroplast.

```

```

FT NON_TER 1 1
FT NON_TER 182 182
SQ SEQUENCE 182 AA; 20945 MW; 6A8E8CA8696D4E80 CRC64;

Query Match
Best Local Similarity 14.3%; Score 7; DB 8; Length 182;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 40 KRPRSGS 46
    |||||
    17 KRPRSGS 23

RESULT 13
O8WL34
ID O8WL34 PRELIMINARY; PRT; 182 AA.
AC O8WL34;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Ribosomal protein S4 (Fragment).
GN RPS4.
OS Gynandriis sisyinchium.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Iridaceae;
OC Gynandriis.
OX NCBI_TaxID=152724;
RN [1]
RP SEQUENCE FROM N.A.
RA Reeves G., Chase M.W., Goldblatt P., Rudall P., Fay M.F., Cox A.V.,
RA Lejeune B., Souza-Chies T.;
RT "Molecular systematics of Iridaceae: Evidence from four plastid DNA
RT regions.";
RL Am. J. Bot. 88:2074-2087(2001).
DR EMBL; AJ409044; CAC83922.1; -.
DR InterPro; IPR001912; Ribosomal_S4.
DR InterPro; IPR002942; S4.
DR InterPro; IPR005709; S4_bact_org.
DR Pfam; PF00163; Ribosomal_S4; 1.
DR Pfam; PF01479; S4; 1.
DR SMART; SM00363; S4; 1.
DR TIGRfams; TIGR01017; tpsd_bact; 1.
DR PROSITE; PS00632; RIBOSOMAL_S4; 1.
KW Chloroplast.
FT NON_TER 1 1
FT NON_TER 182 182
SQ SEQUENCE 182 AA; 20953 MW; 0F70F610315D16C6 CRC64;

Query Match
Best Local Similarity 14.3%; Score 7; DB 8; Length 182;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 40 KRPRSGS 46
    |||||
    17 KRPRSGS 23

RESULT 14
O8WL41
ID O8WL41 PRELIMINARY; PRT; 182 AA.
AC O8WL41;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Ribosomal protein S4 (Fragment).
GN RPS4.
OS Ferraria crispa.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Iridaceae;
OC Ferraria.
OX NCBI_TaxID=148469;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RA Reeves G., Chase M.W., Goldblatt P., Rudall P., Fay M.F., Cox A.V.,
RA Lejeune B., Souza-Chies T.;
RT "Molecular systematics of Iridaceae: Evidence from four plastid DNA
RT regions.";
RL Am. J. Bot. 88:2074-2087(2001).
DR EMBL; AJ409040; CAC83922.1; -.
DR InterPro; IPR001912; Ribosomal_S4.
DR InterPro; IPR002942; S4.
DR InterPro; IPR005709; S4_bact_org.
DR Pfam; PF00163; Ribosomal_S4; 1.
DR Pfam; PF01479; S4; 1.
DR SMART; SM00363; S4; 1.
DR TIGRfams; TIGR01017; tpsd_bact; 1.
DR PROSITE; PS00632; RIBOSOMAL_S4; 1.
KW Chloroplast.
FT NON_TER 1 1
FT NON_TER 182 182
SQ SEQUENCE 182 AA; 21064 MW; 7B5CC415E7B73C30 CRC64;

Query Match
Best Local Similarity 14.3%; Score 7; DB 8; Length 182;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 40 KRPRSGS 46
    |||||
    17 KRPRSGS 23

RESULT 15
O8WL46
ID O8WL46 PRELIMINARY; PRT; 182 AA.
AC O8WL46;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Ribosomal protein S4 (Fragment).
GN RPS4.
OS Diplarrhena latifolia.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Iridaceae;
OC Diplarrhena.
OX NCBI_TaxID=93020;
RN [1]
RP SEQUENCE FROM N.A.
RA Reeves G., Chase M.W., Goldblatt P., Rudall P., Fay M.F., Cox A.V.,
RA Lejeune B., Souza-Chies T.;
RT "Molecular systematics of Iridaceae: Evidence from four plastid DNA
RT regions.";
RL Am. J. Bot. 88:2074-2087(2001).
DR EMBL; AJ409036; CAC83918.1; -.
DR InterPro; IPR001912; Ribosomal_S4.
DR InterPro; IPR002942; S4.
DR InterPro; IPR005709; S4_bact_org.
DR Pfam; PF00163; Ribosomal_S4; 1.
DR Pfam; PF01479; S4; 1.
DR SMART; SM00363; S4; 1.
DR TIGRfams; TIGR01017; tpsd_bact; 1.
DR PROSITE; PS00632; RIBOSOMAL_S4; 1.
KW Chloroplast.
FT NON_TER 1 1
FT NON_TER 182 182
SQ SEQUENCE 182 AA; 20922 MW; ADE02CDB3E240C06 CRC64;

Query Match
Best Local Similarity 14.3%; Score 7; DB 8; Length 182;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 40 KRPRSGS 46
    |||||
    17 KRPRSGS 23

```

Fri Jul 25 15:45:13 2003

us-09-991-681-31.011go-24.rspc

Page 6

Search completed: July 24, 2003, 12:25:54
Job time : 17.0659 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: July 24, 2003, 11:44:42 ; Search time 15.9268 Seconds
(without alignments)
488.334 Million cell updates/sec

Title: US-09-991-681-31

Perfect score: 260

Sequence: 1 EPLPPRGQDSPLLRPQHLM.....SAGPELLRQDRPRSGSTGS 49

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /SIDSI/gcgdata/geneseq/geneeqp-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	260	100.0	49	23	AAO19169 Human prostate-spe
2	260	100.0	49	23	AAO19169 Human prostate-spe
3	260	100.0	518	19	AAO19169 Human prostate-spe
4	260	100.0	518	23	AAO19169 Human prostate-spe
5	260	100.0	518	23	AAO19169 Human prostate-spe
6	260	100.0	1807	22	ABG09728 Human prostate tag
7	260	100.0	1839	23	ABP64835 Human prostate tag
8	260	100.0	1982	22	ABG09731 Human prostate tag
9	67	25.8	195	22	AAU32052 Novel human secret

10	63	24.2	70	22	ABG55616 Human liver peptid
11	63	24.2	70	22	ABG40346 Peptide #7852 enco
12	63	24.2	70	22	AAO19169 Human brain expres
13	63	24.2	70	22	AAO19169 Human bone marrow
14	63	24.2	70	22	AAO19169 Peptide #8084 enco
15	63	24.2	70	23	ABG43753 Human peptide enco
16	63	24.2	171	22	ABG22251 Novel human diagno
17	62	23.8	87	22	AAO19169 Bovine mammary tis
18	60	23.1	1671	23	ABP73759 Candida albicans e
19	59.5	22.9	2004	23	ABG95113 Human translocatio
20	59	22.7	122	22	AAO19169 Propionibacterium
21	58.5	22.5	54	22	AAO19169 Human immune/haema
22	58.5	22.5	3429	22	AAO19169 Drosophila melanog
23	58	22.3	930	24	ABP66278 Human NF-ATc1/C pr
24	57.5	22.1	313	19	AAO19169 Mutated thymidylat
25	57.5	22.1	313	19	AAO19169 Mutated thymidylat
26	57	21.9	71	21	AAO19169 Arabidopsis thalia
27	57	21.9	79	21	AAO19169 Encoded by human G
28	57	21.9	83	13	AAO19169 Human foetal prote
29	57	21.9	152	22	AAO19169 Novel human diagno
30	57	21.9	331	22	ABG27597 Amino acid sequenc
31	57	21.9	485	21	AAO19169 Amino acid sequenc
32	57	21.9	485	21	AAO19169 Human protein SEQ
33	57	21.9	495	22	AAO19169 Human protein SEQ
34	57	21.9	514	22	AAO19169 Novel human diagno
35	57	21.9	558	22	ABG32688 Human thymidylate
36	56.5	21.7	189	22	ABG32688 Mutated thymidylat
37	56.5	21.7	284	23	ABP53743 Mutated thymidylat
38	56.5	21.7	313	19	AAO19169 Mutated thymidylat
39	56.5	21.7	313	19	AAO19169 Mutated thymidylat
40	56.5	21.7	313	19	AAO19169 Mutated thymidylat
41	56.5	21.7	313	19	AAO19169 Mutated thymidylat
42	56.5	21.7	313	19	AAO19169 Mutated thymidylat
43	56.5	21.7	313	19	AAO19169 Mutated thymidylat
44	56.5	21.7	313	19	AAO19169 Mutated thymidylat
45	56.5	21.7	313	19	AAO19169 Mutated thymidylat

ALIGNMENTS

RESULT 1	AAO19169	standard; Protein: 49 AA.
ID	AAO19169	
AC	AAO19169	
XX	27-NOV-2002	(first entry)
DT	27-NOV-2002	
DE	Human prostate-specific PS118 protein fragment #5.	
DE	Human prostate-specific PS118 protein fragment #5.	
XX	Human; prostate; prostate-specific sequence; prostate cancer; PS118;	
KW	EST; expressed sequence tag; cytosolic; gene therapy.	
XX		
OS	Homo sapiens.	
XX		
PN	US2002086316-A1.	
PD	04-JUL-2002.	
XX		
PF	26-NOV-2001; 2001US-0991681.	
XX	23-APR-1998; 98US-0065383.	
PR	23-APR-1997; 97US-0842385.	
XX		
PA	(BILL/) BILLINGEL P A.	
PA	(COHE/) COHEN M.	
PA	(COLP/) COLPITTS T L.	
PA	(FRIE/) FRIEDMAN P N.	
PA	(GORD/) GORDAN J.	
PA	(GRAN/) GRANADOS E N.	
PA	(HODG/) HODGES S C.	
PA	(KLAS/) KLAS M R.	

PA (KRAT/) KRATOCHVIL J D.
 PA (ROBE/) ROBERTS-RAPP L.
 PA (RUSS/) RUSSELL J C.
 PA (STRO/) STROUPE S D.
 PI Billingsel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;
 PI Granados EN, Hodges SC, Klasse MR, Kratochvil JD, Roberts-Rapp L;
 PI Russell JC, Stroupe SD;
 XX WPI: 2002-665429/71.
 XX
 PT Novel PS18 polypeptide for detecting, diagnosing, staging, monitoring,
 PT prognosticating, preventing, treating, or determining predisposition of
 PT individual to diseases and conditions of prostate, e.g. prostate
 PT cancer -
 PS Claim 17; Page 44; 58pp; English.
 CC The present invention relates to a number of prostate-specific sequences
 CC derived from the human PS18 gene. These can be used in the detection,
 CC monitoring and treatment of prostate diseases, particularly prostate
 CC cancer. The present sequence is a PS18 protein fragment of the
 CC invention. The coding sequences of the invention were isolated from a
 CC prostate tissue expressed sequence tag (EST) library.
 CC
 SQ Sequence 49 AA;
 Query Match 100.0%; Score 260; DB 23; Length 49;
 Best Local Similarity 100.0%; Pred. No. 1.5e-28;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EPLGRGQDSPLLRPQHLMDOGMRHSFSAGPELLRQDKRPRSGSTGS 49
 Db 1 EPLGRGQDSPLLRPQHLMDOGMRHSFSAGPELLRQDKRPRSGSTGS 49
 RESULT 2
 ID AAM50813 standard; Protein; 49 AA.
 AC AAM50813;
 XX
 DT 01-MAY-2002 (first entry)
 DE PS18 prostate marker immunogenic polypeptide.
 XX
 KW PS18; prostate; marker; prostate cancer; tumour; metastasis;
 KW benign prostatic hyperplasia; prostatic intraepithelial neoplasia;
 KW prostatictis; human; diagnosis; therapy; vaccine; immunogen.
 XX
 OS Homo sapiens.
 XX
 PN US2001055758-A1.
 PD 27-DEC-2001.
 XX
 PF 23-APR-1998; 98US-0065383.
 PR 23-APR-1997; 97US-0842385.
 XX
 PA (BILL/) BILLING-MEDEL P A.
 PA (COHE/) COHEN M.
 PA (COPL/) COPLITTIS T L.
 PA (FRIE/) FRIEDMAN P N.
 PA (GORD/) GORDON J.
 PA (GRAN/) GRANADOS E N.
 PA (HODG/) HODGES S C.
 PA (KLAS/) KLAS M R.
 PA (KRAT/) KRATOCHVIL J D.
 PA (ROBE/) ROBERTS-RAPP L.
 PI Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;
 PI Granados EN, Hodges SC, Klasse MR, Kratochvil JD, Roberts-Rapp L;

PI Russell JC, Stroupe SD;
 XX
 DR WPI: 2002-187683/24.
 XX
 PT Detecting presence of target PS18 polynucleotide in test sample,
 PT useful for detecting, diagnosing, staging, monitoring, prognosticating,
 PT preventing or treating or determining predisposition to prostate
 PT disease -
 PS Claim 17; Page 44; 57pp; English.
 CC The present sequence is that of an immunogenic polypeptide
 CC comprising amino acids 393-441 of human prostate-specific PS18
 CC polypeptide (see AAM50809). A PS18 consensus sequence (see
 CC ABA91651) is found at least 12 times more often in prostate than
 CC in non-prostate tissue. PS18 polypeptides, polynucleotides,
 CC antibodies, agonists and inhibitors are useful for detecting,
 CC diagnosing, staging, monitoring, prognosticating, preventing and
 CC treating (including by genetic immunisation), or determining the
 CC predisposition of an individual to, diseases and conditions of the
 CC prostate, such as benign prostatic hyperplasia, prostaticitis,
 CC prostatic intraepithelial neoplasia, prostate cancer, tumours and
 CC metastases. The PS18 polypeptides can be produced by expression
 CC of PS18 polynucleotides in transfected host cells, and
 CC immunogenic peptides are useful for raising PS18-specific
 CC antibodies of diagnostic use. The methods and reagents of the
 CC invention may provide an early means of detecting diseases of the
 CC prostate and may also provide new markers which can differentiate
 CC between the clinically important and unimportant prostate cancers
 CC without the use of surgery.
 CC
 SQ Sequence 49 AA;
 Query Match 100.0%; Score 260; DB 23; Length 49;
 Best Local Similarity 100.0%; Pred. No. 1.5e-28;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EPLGRGQDSPLLRPQHLMDOGMRHSFSAGPELLRQDKRPRSGSTGS 49
 Db 1 EPLGRGQDSPLLRPQHLMDOGMRHSFSAGPELLRQDKRPRSGSTGS 49
 RESULT 3
 ID AAM65472 standard; Protein; 518 AA.
 AC AAM65472;
 XX
 DT 25-MAR-2003 (updated)
 DT 25-FEB-1999 (first entry)
 DE PS18 protein encoded by consensus sequence.
 XX
 KW EST clone; PS18; prostate tumour tissue; prostatic disease; cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO9848054-A1.
 PD 29-OCT-1998.
 XX
 PF 23-APR-1998; 98WO-US08239.
 PR 23-APR-1997; 97US-0842385.
 XX
 PA (ABBO) ABBOTT LAB.
 PI Billing-medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;
 PI Granados EN, Hodges SC, Klasse MR, Kratochvil JD, Robertarapp L;
 PI Russell JC, Stroupe SD;
 XX
 DR WPI: 1998-610000/51.
 DR N-PSDB; AAV82812.

XX New P118 nucleic acid and proteins - used for diagnosis and
PT treatment of prostatic disease, especially cancer, and also for drug
PT screening
PS Claim 17; Page 93-94; 117pp; English.
XX
XX The present sequence is encoded by consensus P5118 sequence derived from
CC overlapping EST (expressed sequence tag) clones AAV82803-10. The clones
CC were identified from a cDNA library made from prostate tumour tissue.
CC Recombinant P5118 protein is used to detect P5118-specific antibodies,
CC to raise antibodies for detection of P5118 antigens, to screen for
CC specific binding agents (potential therapeutics), and to isolate specific
CC antibodies from serum. Detection of P5118 protein or nucleic acid, which
CC are prostate related, and altered or elevated in prostatic disease, is
CC used for detection, diagnosis, staging, monitoring and prognosis of
CC prostatic disease, particularly cancer, and to identify subjects at
CC risk.
CC (Updated on 25-MAR-2003 to correct PI field.)
CC
SQ Sequence 518 AA;
Query Match 100.0%; Score 260; DB 19; Length 518;
Best Local Similarity 100.0%; Pred. No. 2.4e-27;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EPLGPRGDSPLLRPOHLMDOGQWRHSFSAAGPELLRQDKRPRSGTGS 49
DB 393 EPLGPRGDSPLLRPOHLMDOGQWRHSFSAAGPELLRQDKRPRSGTGS 441
RESULT 4
ID AAO19165 standard; Protein; 518 AA.
XX
XX AAO19165;
XX
XX 27-NOV-2002 (first entry)
XX
XX Human prostate-specific P5118 protein fragment #1.
XX
XX Human; prostate; prostate-specific sequence; prostate cancer; P5118;
XX EST; expressed sequence tag; cytostatic; gene therapy.
XX
XX Homo sapiens.
XX
XX US2002086316-A1.
XX
XX 04-JUL-2002.
XX
XX 26-NOV-2001; 2001US-0991681.
XX
XX 23-APR-1998; 98US-0065383.
XX
XX 23-APR-1997; 97US-0842385.
XX
XX (BILL/) BILLINGEL P A.
XX (COHE/) COHEN M.
XX (COPL/) COPLITTIS T L.
XX (FRIE/) FRIEDMAN P N.
XX (GORD/) GORDAN J.
XX (GRAN/) GRANADOS E N.
XX (HODG/) HODGES S C.
XX (KLAS/) KLAAS M R.
XX (KRAT/) KRATOCHVIL J D.
XX (ROBE/) ROBERTS-RAPP L.
XX (RUSSE/) RUSSELL J C.
XX (STRO/) STROUPE S D.
XX
XX Billingel PA, Cohen M, Coplittis TL, Friedman PN, Gordon J;
PI Granados EN, Hodges SC, Klaas MR, Kratochvil JD, Roberts-Rapp L;
PI Russell JC, Stroupe SD;
XX WPI; 2002-665429/71.
DR

XX Novel P5118 polypeptide for detecting, diagnosing, staging, monitoring,
PT prognosticating, preventing, treating, or determining predisposition of
PT individual to diseases and conditions of prostate, e.g. prostate
PT cancer
PS Claim 17; Page 42-43; 58pp; English.
XX
XX The present invention relates to a number of prostate-specific sequences
CC derived from the human P5118 gene. These can be used in the detection,
CC monitoring and treatment of prostate diseases, particularly prostate
CC cancer. The present sequence is a P5118 protein fragment of the
CC invention. The coding sequences of the invention were isolated from a
CC prostate tissue expressed sequence tag (EST) library.
XX
SQ Sequence 518 AA;
Query Match 100.0%; Score 260; DB 23; Length 518;
Best Local Similarity 100.0%; Pred. No. 2.4e-27;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EPLGPRGDSPLLRPOHLMDOGQWRHSFSAAGPELLRQDKRPRSGTGS 49
DB 393 EPLGPRGDSPLLRPOHLMDOGQWRHSFSAAGPELLRQDKRPRSGTGS 441
RESULT 5
ID AAM50809 standard; Protein; 518 AA.
XX
XX AAM50809;
XX
XX 01-MAY-2002 (first entry)
XX
XX P5118 prostate marker partial sequence.
XX
XX P5118; prostate; marker; prostate cancer; tumour; metastasis;
XX benign prostatic hyperplasia; prostatic intraepithelial neoplasia;
XX prostatitis; human; diagnosis; therapy; vaccine.
XX
XX Homo sapiens.
XX
XX US2001055758-A1.
XX
XX 27-DEC-2001.
XX
XX 23-APR-1998; 98US-0065383.
XX
XX 23-APR-1997; 97US-0842385.
XX
XX (BILL/) BILLING-MEDEL P A.
XX (COHE/) COHEN M.
XX (COPL/) COPLITTIS T L.
XX (FRIE/) FRIEDMAN P N.
XX (GORD/) GORDON J.
XX (GRAN/) GRANADOS E N.
XX (HODG/) HODGES S C.
XX (KLAS/) KLAAS M R.
XX (KRAT/) KRATOCHVIL J D.
XX (ROBE/) ROBERTS-RAPP L.
XX
XX Billing-Medel PA, Cohen M, Coplittis TL, Friedman PN, Gordon J;
PI Granados EN, Hodges SC, Klaas MR, Kratochvil JD, Roberts-Rapp L;
PI Russell JC, Stroupe SD;
XX WPI; 2002-187683/24.
XX N-PSDB; ABA91651.
XX
XX Detecting presence of target P5118 polynucleotide in test sample,
PT useful for detecting, diagnosing, staging, monitoring, prognosticating,
PT preventing or treating or determining predisposition to prostate
PT disease -
XX

PS Claim 17; Page 42-43; 57pp; English.

CC The present sequence is that of a human prostate-specific PS118
 CC polypeptide, as predicted from a partial consensus cDNA sequence
 CC (see AB91651), and lacking the N-terminal region. The PS118
 CC consensus sequence is found at least 12 times more often in
 CC prostate than in non-prostate tissue. PS118 polypeptides,
 CC including derivatives of the present sequence, polynucleotides,
 CC antibodies, agonists and inhibitors are useful for detecting,
 CC diagnosing, staging, monitoring, prognosticating, preventing and
 CC treating (including by genetic immunisation), or determining the
 CC predisposition of an individual to, diseases and conditions of the
 CC prostate, such as benign prostatic hyperplasia, prostatitis,
 CC prostatic intraepithelial neoplasia, prostate cancer, tumours and
 CC metastases. The PS118 polypeptides can be produced by expression
 CC of PS118 polynucleotides in transfected host cells. The methods
 CC and reagents of the invention may provide an early means of
 CC detecting diseases of the prostate and may also provide new markers
 CC which can differentiate between the clinically important and
 CC unimportant prostate cancers without the use of surgery.

XX SQ Sequence 518 AA;

QY Query Match 100.0%; Score 260; DB 23; Length 518;
 Best Local Similarity 100.0%; Pred. No. 2,4e-27;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 EPLGRGDSPLLRPPHLMDOGMRHSFSAPELLRQDKRRSGSTGS 49
 |||||
 393 EPLGRGDSPLLRPPHLMDOGMRHSFSAPELLRQDKRRSGSTGS 441

RESULT 6
 ABC09728
 ID ABG09728 standard; Protein; 1807 AA.

XX AC ABC09728;

XX DT 13-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #9719.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Dymnac RT, Liu C, Tang YT;

XX DR WPI: 2001-639362/73.

XX DR N-PSDB; AAS73915.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -

XX PS Claim 20; SEQ ID No 40087; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 1807 AA;

QY Query Match 100.0%; Score 260; DB 22; Length 1807;
 Best Local Similarity 100.0%; Pred. No. 1.1e-26;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 EPLGRGDSPLLRPPHLMDOGMRHSFSAPELLRQDKRRSGSTGS 49
 |||||
 1682 EPLGRGDSPLLRPPHLMDOGMRHSFSAPELLRQDKRRSGSTGS 1730

RESULT 7
 ABP64835
 ID ABP64835 standard; Protein; 1839 AA.

XX AC ABP64835;

XX DT 25-FEB-2003 (first entry)

XX DE Human protein SEQ ID 495.

XX KW Human; expressed sequence tag; EST;
 KW haematopoietic disorder; central nervous system disease; viral infection;
 KW peripheral nervous system disease; non-healing wound; infectious disease;
 KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;
 KW fungal infection; autoimmune disorder; coagulation disorder; noctropic;
 KW anti-allergic; anti-inflammatory; immunosuppressive; neuroprotective;
 KW cytostatic; haemostatic; virucide; antibacterial; fungicide;
 KW immunostimulant; cerebroprotective.

XX OS Homo sapiens.

XX PN WO200259260-A2.

XX PD 01-AUG-2002.

XX PF 16-NOV-2001; 2001WO-US42950.

XX PR 17-NOV-2000; 2000US-0714936.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;
 PI Ren F, Xue AJ, Yang Y, Wehrman T, Dymnac RT;

XX DR WPI: 2002-590824/63.

XX DR N-PSDB; ABQ99421.

XX PT New isolated polynucleotide, useful in research, diagnostic or
 PT therapeutic methods, e.g. preventing or treating disorders involving
 PT aberrant protein expression or biological activity -

XX PS Claim 20; SEQ ID 495; 394pp; English.

CC The present invention relates to novel human coding sequences
CC (ABG09268-ABG09608) and proteins (ABP64682-ABP65022). The sequences are
CC useful in therapeutic, diagnostic and research methods. The
CC polynucleotides may be used in the field of molecular biology as
CC hybridisation probes, primers for PCR, for chromosome and gene mapping,
CC for the recombinant production of protein, or in generation of anti-sense
CC DNA or RNA. The polynucleotides are useful in diagnostics as expressed
CC sequence tags (ESTs) for identifying expressed genes or for physical
CC mapping of the human genome. The proteins may be used as molecular weight
CC markers, or as nutritional sources or supplements. The proteins may be
CC used to maintain and expand cell population in a totipotent or
CC pluripotent state useful for re-engineering damaged or diseased
CC tissues, transplantation, manufacture of bio-pharmaceuticals or the
CC development of bio-sensors. The polynucleotides and proteins are useful
CC for preventing, treating or ameliorating disorders involving aberrant
CC protein expression or biological activity, e.g. haematopoietic disorders,
CC central/peripheral nervous system diseases, mechanical and traumatic
CC disorders, non-healing wounds, immune deficiencies and disorders,
CC infectious diseases caused by viral, bacterial or fungal infection,
CC autoimmune disorders, allergic reactions and conditions, coagulation
CC disorders, or cancer. The polynucleotide sequences of the invention were
CC assembled from ESTs isolated mainly by sequencing by hybridisation, and
CC in some cases, sequences obtained from one or more public databases.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 1839 AA;

Query Match 100.0%; Score 260; DB 23; Length 1839;
Best Local Similarity 100.0%; Pred. No. 1.1e-26;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EPILPRGQDSPLLRPQHLMQGMKHSFSGAPELLRODKRPRSGSTGS 49
Db 1714 EPILPRGQDSPLLRPQHLMQGMKHSFSGAPELLRODKRPRSGSTGS 1762

RESULT 8
ABG09731
ID ABG09731 standard; Protein; 1982 AA.

XX ABG09731;
AC
XX
XX
DT 13-FEB-2002 (first entry)
XX

DE Novel human diagnostic protein #9722.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX

OS Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; NAST73918.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

XX Claim 20; SEQ ID No 40090; 103bp; English.

XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. ABG00010-ABG30377 represent novel human
XX diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 1982 AA;

Query Match 100.0%; Score 260; DB 22; Length 1982;
Best Local Similarity 100.0%; Pred. No. 1.2e-26;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EPILPRGQDSPLLRPQHLMQGMKHSFSGAPELLRODKRPRSGSTGS 49
Db 1796 EPILPRGQDSPLLRPQHLMQGMKHSFSGAPELLRODKRPRSGSTGS 1844

RESULT 9
AAU32052
ID AAU32052 standard; Protein; 195 AA.

XX AAU32052;
AC
XX
XX
DT 18-DEC-2001 (first entry)
XX

DE Novel human secreted protein #2543.

XX Human; vaccination; gene therapy; nutritional supplement;

KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

XX Homo sapiens.

XX WO200179449-A2.

XX 25-OCT-2001.

XX 16-APR-2001; 2001WO-US08656.

XX 18-APR-2000; 2000US-0552929.

XX 26-JAN-2001; 2001US-0770160.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-611725/70.

XX Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy -

XX Claim 20; Page 548-549; 765pp; English.

XX The invention relates to novel human secreted polypeptides. The

CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising
 CC the nucleic acids encoding the polypeptides and cells genetically
 CC engineered to express them are also useful for producing the proteins.
 CC The proteins are useful in genetic vaccination, testing and
 CC therapy, and can be used as nutritional supplements. They may be used to
 CC increase stem cell proliferation; to regulate haemopoiesis; and in
 CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
 CC immune suppression and/or stimulation; as anti-inflammatory agents; and
 CC in treatment of leukaemias. AAU29510-AAU3304 represent the amino acid
 CC sequences of novel human secreted proteins of the invention.

CC Sequence 195 AA;

Query Match 25.8%; Score 67; DB 22; Length 195;
 Best Local Similarity 37.5%; Pred. No. 0.44;
 Matches 18; Conservative 6; Mismatches 18; Indels 6; Gaps 2;

Oy 4 GPRGDSPLLRPOHLMDOGMHSPFSAAGPELLRQDK--RPRSGSTGS 49
 ID ABB40346 standard; Peptide; 70 AA.
 Db 55 GKNQGGEPVAVVHSCSHL---VKHSQSRPSSWRQEKTRITVGGGPGS 98

RESULT 10

ABG55616
 ID ABB55616 standard; Peptide; 70 AA.

AC ABB55616;

DT 25-FEB-2003 (first entry)

DE Human liver peptide, SEQ ID No 34264.

KM Human, liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;

KW hypercholesterolaemia; coronary heart disease.

OS Homo sapiens.

PN WO200157273-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00664.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0609408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-488898/53.

PT Human genome-derived single exon nucleic acid probes useful for

CC analysing gene expression in human adult liver -

PS Claim 27; SEQ ID No 34264; 658bp; English.

CC The invention relates to a single exon nucleic acid probe (SENP) (1) for

CC measuring human gene expression in a sample derived from human adult

CC liver, comprising one of 13109 defined nucleotide sequences given in the

CC specification (or complements/ fragments). The probe hybridises at high

CC liver. (1) may be used for predicting, measuring and displaying gene
 CC expression in samples derived from human adult liver. The genes
 CC identified may be involved in genetic liver diseases such as cirrhosis,
 CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which
 CC is associated with coronary heart disease. ABB47348-ABG59930 represent
 CC human liver single exon encoded peptides of the invention.
 CC Note: The sequence information for this patent does not appear in the
 CC printed specification but was obtained in electronic format directly
 CC from WIPO at http://wipo.int/pub/published_pct_sequences.

CC Sequence 70 AA;

Query Match 24.2%; Score 63; DB 22; Length 70;
 Best Local Similarity 34.5%; Pred. No. 0.48;
 Matches 19; Conservative 7; Mismatches 13; Indels 16; Gaps 3;

Oy 5 PRGQDSPLLRPOHLMDOGMHSPFSA-----GP-----ELLRQDKRPRSGSTG 48
 ID ABB40346
 Db 12 PRGQ-----QDRPHLHHRGGRHSEAVVSHRPGQVRIHLQAVRDEAGLGAGTG 61

RESULT 11

ABB40346
 ID ABB40346 standard; Peptide; 70 AA.

AC ABB40346;

DT 04-FEB-2002 (first entry)

DE Peptide #7852 encoded by human foetal liver single exon probe.

KM Human, foetal liver; gene expression; single exon nucleic acid probe.

OS Homo sapiens.

PN WO200157273-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00669.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0609408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-483447/52.

PT Human genome-derived single exon nucleic acid probes useful for

CC analysing gene expression in human fetal liver -

PS Claim 27; SEQ ID NO 32981; 639bp + sequence listing; English.

CC The invention relates to a single exon nucleic acid probe for

CC measuring human gene expression in a sample derived from human foetal

CC liver. The single exon nucleic acid probes may be used for predicting,

CC measuring and displaying gene expression in samples derived from human

CC fetal liver. The present sequence is a peptide encoded by a single exon

CC nucleic acid probe of the invention.

CC Note: The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly

CC from WIPO at http://wipo.int/pub/published_pct_sequences.

CC Sequence 70 AA;

Query Match 24.2%; Score 63; DB 22; Length 70;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 11:53:53 ; Search time 4.37628 Seconds
(without alignments)
473.743 Million cell updates/sec

Title: US-09-991-681-31

Perfect score: 260
Sequence: 1 EPLGPRGQDSEPLRQPHLM.....SAGPELLRQDRKPRSGSTGS 49

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

1: /cgn2_6/ptodata/1/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTUS.COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	260	100.0	49	US-09-065-383-31	Sequence 31, Appl
2	260	100.0	518	US-09-065-383-27	Sequence 27, Appl
3	70.5	27.1	309	US-09-252-991A-17975	Sequence 17975, A
4	65	25.0	693	US-09-252-991A-24059	Sequence 24059, A
5	64.5	24.8	501	US-09-252-991A-21596	Sequence 21596, A
6	63	24.2	504	US-09-252-991A-23374	Sequence 23374, A
7	62	23.8	567	US-09-252-991A-28435	Sequence 28435, A
8	61	23.5	171	US-09-252-991A-18380	Sequence 18380, A
9	59	22.7	109	US-09-252-991A-28987	Sequence 28987, A
10	58.5	22.5	681	US-09-252-991A-22519	Sequence 22519, A
11	58	22.3	297	US-09-252-991A-29217	Sequence 29217, A
12	57.5	22.1	140	US-09-252-991A-28444	Sequence 28444, A
13	57.5	22.1	148	US-09-252-991A-26802	Sequence 26802, A
14	57	21.9	351	US-09-252-991A-17299	Sequence 17299, A
15	57	21.9	565	US-09-252-991A-28571	Sequence 28571, A
16	57	21.9	566	US-09-252-991A-32371	Sequence 32371, A
17	57	21.9	592	US-09-252-991A-18124	Sequence 18124, A
18	56.5	21.7	312	US-09-347-878-30	Sequence 34, Appl
19	56.5	21.7	313	US-09-367-007C-39	Sequence 30, Appl
20	56.5	21.7	313	US-09-367-007C-39	Sequence 39, Appl
21	56.5	21.7	558	US-09-252-991A-26115	Sequence 26115, A
22	56	21.5	163	US-09-066-074-2	Sequence 2, Appl
23	56	21.5	163	US-08-555-912A-2	Sequence 2, Appl
24	56	21.5	163	US-09-208-804-4	Sequence 4, Appl
25	56	21.5	163	US-08-801-743-4	Sequence 4, Appl
26	56	21.5	239	US-09-252-991A-21250	Sequence 21250, A
27	56	21.5	429	US-09-252-991A-16841	Sequence 16841, A

28	56	21.5	536	US-09-252-991A-23495	Sequence 23495, A
29	56	21.5	541	US-09-252-991A-17206	Sequence 17206, A
30	56	21.5	620	US-09-252-991A-17304	Sequence 17304, A
31	55.5	21.3	199	US-09-252-991A-23002	Sequence 23002, A
32	55.5	21.3	472	US-09-252-991A-30367	Sequence 30367, A
33	55	21.2	448	US-09-252-991A-24066	Sequence 24066, A
34	55	21.2	516	US-09-252-991A-31707	Sequence 31707, A
35	55	21.2	1882	US-09-369-364A-13	Sequence 13, Appl
36	54.5	21.0	336	US-09-252-991A-20600	Sequence 20600, A
37	54.5	21.0	365	US-09-252-991A-26578	Sequence 26578, A
38	54.5	21.0	647	US-09-031-563-7	Sequence 7, Appl
39	54.5	21.0	647	US-09-392-277-7	Sequence 7, Appl
40	54.5	21.0	647	US-09-258-000-7	Sequence 7, Appl
41	54.5	21.0	648	US-09-031-563-5	Sequence 5, Appl
42	54.5	21.0	648	US-09-392-277-5	Sequence 5, Appl
43	54.5	21.0	648	US-09-258-000-5	Sequence 5, Appl
44	54.5	21.0	1055	US-09-031-563-27	Sequence 27, Appl
45	54.5	21.0	1055	US-09-392-277-27	Sequence 27, Appl

ALIGNMENTS

RESULT 1
US-09-065-383-31
Sequence 31, Application US/09065383
Patent No. 6391543
GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GORDON, JULIAN
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KLASS, MICHAEL R.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: ROBERTS-RAPP, LISA
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
TITLE OF INVENTION: FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/065,383
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/842,385
FILING DATE: 23-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6084.US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6391543e
US-09-065-383-31

Query Match 100.0%; Score 260; DB 4; Length 49;
Best Local Similarity 100.0%; Pred. No. 6.5e-30;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EPLGRGDSPLLRPQHLMDOGMHFSAGPELLRQDKRRSGSTGS 49
Db 1 EPLGRGDSPLLRPQHLMDOGMHFSAGPELLRQDKRRSGSTGS 49

RESULT 2
US-09-065-383-27

Sequence 27, Application US/09065383
Patent No. 6391543

GENERAL INFORMATION:

APPLICANT: BILLING-MEDEL, PATRICIA

APPLICANT: COHEN, MAURICE

APPLICANT: COLPITTS, TRACEY L.

APPLICANT: FRIEDMAN, PAULA N.

APPLICANT: GORDON, JULIAN

APPLICANT: GRANADOS, EDWARD N.

APPLICANT: HODGES, STEVEN C.

APPLICANT: KLAAS, MICHAEL R.

APPLICANT: KRATOCHVIL, JON D.

APPLICANT: ROBERTS-RAPP, LISA

APPLICANT: RUSSELL, JOHN C.

APPLICANT: STROUPE, STEPHEN D.

TITLE OF INVENTION: REAGENTS AND METHODS USEFUL

NUMBER OF INVENTION: FOR DETECTING DISEASES OF THE PROSTATE

NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESS:

ADDRESSEE: Abbott Laboratories

STREET: 100 Abbott Park Road

CITY: Abbott Park

STATE: IL

COUNTRY: USA

ZIP: 60064-3500

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/065,383

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/842,385

FILING DATE: 23-APR-1997

ATTORNEY/AGENT INFORMATION:

NAME: Becker, Cheryl L.

REGISTRATION NUMBER: 35,441

REFERENCE/DOCKET NUMBER: 6084.US.P1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 847/935-1729

TELEFAX: 847/938-2623

TELEX:

INFORMATION FOR SEQ ID NO: 27:

SEQUENCE CHARACTERISTICS:

LENGTH: 518 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: No. 6391543e

US-09-065-383-27

Query Match 100.0%; Score 260; DB 4; Length 518;
Best Local Similarity 100.0%; Pred. No. 1.1e-28;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EPLGRGDSPLLRPQHLMDOGMHFSAGPELLRQDKRRSGSTGS 49
Db 393 EPLGRGDSPLLRPQHLMDOGMHFSAGPELLRQDKRRSGSTGS 441

RESULT 3
US-09-252-991A-17975

Sequence 17975, Application US/09252991A
Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A

CURRENT FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 17975

LENGTH: 309

TYPE: PRT

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-17975

Query Match 27.1%; Score 70.5; DB 4; Length 309;
Best Local Similarity 41.5%; Pred. No. 0.043;
Matches 22; Conservative 3; Mismatches 19; Indels 9; Gaps 2;

Qy 2 PLGRGDSPLLRPQHLMDOGMHFSAGPELLRQDKRRSGSTGS 45
Db 194 PLGRGDSPLLRPQHLMDOGMHFSAGPELLRQDKRRSGSTGS 246

RESULT 4
US-09-252-991A-24059

Sequence 24059, Application US/09252991A
Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A

CURRENT FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 24059

LENGTH: 693

TYPE: PRT

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-24059

Query Match 25.0%; Score 65; DB 4; Length 693;
Best Local Similarity 29.2%; Pred. No. 0.7;
Matches 19; Conservative 5; Mismatches 19; Indels 22; Gaps 1;

Qy 2 PLGRGDSPLLRPQHLMDOGMHFSAGPELLRQDKRRSGSTGS 39
Db 310 PLGRGDSPLLRPQHLMDOGMHFSAGPELLRQDKRRSGSTGS 369

Qy 40 KRPRS 44
Db 370 HRPPS 374


```
RESULT 5
US-09-252-991A-21596
; Sequence 21596, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21596
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21596

Query Match          24.8%; Score 64.5; DB 4; Length 501;
Best Local Similarity 33.3%; Pred. No. 0.55;
Matches 15; Conservative 5; Mismatches 14; Indels 11; Gaps 1;

OY      2  PLGPRGDSPLLRPOHLMDOGMHFSAGPELRLRDKRPRS 46
DB      368  PVGPGDQDGP-----AAQAGRGFRPDRRLRGRGRLPGA 401

RESULT 6
US-09-252-991A-23374
; Sequence 23374, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23374
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23374

Query Match          24.2%; Score 63; DB 4; Length 504;
Best Local Similarity 48.3%; Pred. No. 0.91;
Matches 14; Conservative 3; Mismatches 10; Indels 2; Gaps 1;

OY      2  PLGPRGDSPLLRPOHLMDOGMHFS 30
DB      364  PLRERGEDIPILFR--HFAEAGAMRHGLT 390

RESULT 7
US-09-252-991A-28435
; Sequence 28435, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28435
; LENGTH: 567
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28435

Query Match          23.8%; Score 62; DB 4; Length 567;
Best Local Similarity 37.2%; Pred. No. 1.5;
Matches 16; Conservative 4; Mismatches 21; Indels 2; Gaps 1;

OY      1  EPGLPRGDSPLLRPOHLMDOGMHFSAGPELRLRDKRPR 43
DB      156  QPTGPAAGARRQHPQHRLDVQHLR--FRPPGLRQGGHR 196

RESULT 8
US-09-252-991A-18380
; Sequence 18380, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18380
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18380

Query Match          23.5%; Score 61; DB 4; Length 171;
Best Local Similarity 64.7%; Pred. No. 0.47;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY      4  GPRGDSPLLRPOHLM 20
DB      142  GPRGKNSPWLQTPSHM 158

RESULT 9
US-09-252-991A-28987
; Sequence 28987, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28987
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28987
```

```
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28435
; LENGTH: 567
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28435

Query Match          23.8%; Score 62; DB 4; Length 567;
Best Local Similarity 37.2%; Pred. No. 1.5;
Matches 16; Conservative 4; Mismatches 21; Indels 2; Gaps 1;

OY      1  EPGLPRGDSPLLRPOHLMDOGMHFSAGPELRLRDKRPR 43
DB      156  QPTGPAAGARRQHPQHRLDVQHLR--FRPPGLRQGGHR 196

RESULT 8
US-09-252-991A-18380
; Sequence 18380, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18380
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18380

Query Match          23.5%; Score 61; DB 4; Length 171;
Best Local Similarity 64.7%; Pred. No. 0.47;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY      4  GPRGDSPLLRPOHLM 20
DB      142  GPRGKNSPWLQTPSHM 158

RESULT 9
US-09-252-991A-28987
; Sequence 28987, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28987
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28987
```

Query Match 22.7%; Score 59; DB 4; Length 109;
 Best Local Similarity 33.3%; Pred. No. 0.52;
 Matches 15; Conservative 8; Mismatches 16; Indels 6; Gaps 2;

Qy 2 PLGPRGDSPLLRPOHLMDOGMHRSFAGPELLRQDKRP-RSG 45
 Db 68 PAFGSRVADVPRSPRR-----ARRRGIAAGRAAPHRBSRPVRSRG 107

RESULT 10
 US-09-252-991A-22519

; Sequence 22519, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 22519

; LENGTH: 681

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-22519

Query Match 22.5%; Score 58.5; DB 4; Length 681;
 Best Local Similarity 33.3%; Pred. No. 5.7;
 Matches 14; Conservative 4; Mismatches 15; Indels 9; Gaps 1;

Qy 2 PLGPRGDSPLLRPOHLMDOGMHRSFAGPELLRQDKRP 43
 Db 237 PAFGPRARQHPRLOQPLPADHAHVRH-----RQARRPR 269

RESULT 11

US-09-252-991A-29217

; Sequence 29217, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 29217

; LENGTH: 297

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-29217

Query Match 22.3%; Score 58; DB 4; Length 297;
 Best Local Similarity 38.0%; Pred. No. 2.4;
 Matches 19; Conservative 6; Mismatches 19; Indels 6; Gaps 3;

Qy 1 EPLGPRGDSPLLRPOHLMDOGMHRSFAGPELLRQDKRP-RSG 48
 Db 213 QFGDPFGQ--PLQGRQPD--PQGRHLVLRARRLGRRAEGRPAGAHG 258

RESULT 12

US-09-252-991A-28444

; Sequence 28444, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 28444

; LENGTH: 140

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-28444

Query Match 22.1%; Score 57.5; DB 4; Length 140;
 Best Local Similarity 36.4%; Pred. No. 1.2;
 Matches 16; Conservative 5; Mismatches 20; Indels 3; Gaps 2;

Qy 1 EPLGPR--GDSPLLRPOHLMDOGMHRSFAGPELLRQDKRP 42
 Db 21 QPABRVGGQPPARVFRDRRRRQGO--RHPDRALRALDRPRP 63

RESULT 13

US-09-252-991A-26802

; Sequence 26802, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 26802

; LENGTH: 148

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-26802

Query Match 22.1%; Score 57.5; DB 4; Length 148;
 Best Local Similarity 29.6%; Pred. No. 1.2;
 Matches 21; Conservative 5; Mismatches 18; Indels 27; Gaps 3;

Qy 5 PRGDSPLLRPOHLMDOGMHRSFAGPELLRQDKRP 37
 Db 39 PGLHPLRLRGQGPAPAPGRPHRTDQARRRHSGADHDQLRQPAFGGPFDEGRGLHR 98

Qy 38 QDKRPSGSGTG 48
 Db 99 QAFRRRNAPG 109

RESULT 14

US-09-252-991A-17299

; Sequence 17299, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

```

1 CURRENT APPLICATION NUMBER: US/09/252,991A
2 CURRENT FILING DATE: 1999-02-18
3 PRIOR APPLICATION NUMBER: US 60/074,788
4 PRIOR FILING DATE: 1998-02-18
5 PRIOR APPLICATION NUMBER: US 60/094,190
6 PRIOR FILING DATE: 1998-07-27
7 NUMBER OF SEQ ID NOS: 31142
8 SEQ ID NO 17299
9 LENGTH: 351
10 TYPE: PR1
11 ORGANISM: Pseudomonas aeruginosa
12 US-09-252-991A-17299

```

Query Match	21.9%;	Score 57;	DB 4;	Length 351;
Best Local Similarity	38.5%;	Pred. No. 4.2;		
Matches	15;	Conservative	8;	Mismatches 8; Indels 8; Gaps 2

```
QY      2 PLGPRGQDSPLLQRPQHLMDOGMRHSFSAGPELLRODK 40
          |||::: |||::: |||:::
Db      236 PVGGRGEEER---QRHRLPDRGR-----RADPPGLRQER 266
```

```

RESULT 15
US-09-252-991A-28571
: Sequence 28571, Application US/09252991A
: Patent No. 6551795
: GENERAL INFORMATION:
: APPLICANT: Marc J. Rubenfield et al.
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
: TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: 107196.136
: CURRENT APPLICATION NUMBER: US/09/252,991A
: CURRENT FILING DATE: 1999-02-18
: PRIOR APPLICATION NUMBER: US 60/074,788
: PRIOR FILING DATE: 1998-02-18
: PRIOR APPLICATION NUMBER: US 60/094,190
: PRIOR FILING DATE: 1998-07-27
: NUMBER OF SEQ ID NOS: 33142
: SEQ ID NO 28571
: LENGTH: 565
: TYPE: PRT
: ORGANISM: Pseudomonas aeruginosa
: US-09-252-991A-28571

```

Query Match	21.9%;	Score 57;	DB 4;	Length 565;
Best Local Similarity	36.1%;	Pred. No. 7.4;		
Matches 13;	Conservative 8;	Mismatches 13;	Indels 2;	Gaps 1

QY 13 LQRPRLHLDGQGRHNSFSAGPELLRQDKPRSGSTG 48
|||:::|:|||||::|
Db 243 LQAPREIYE-KRGRYDAGPEALRQERLARLKEIG 276

Search completed: July 24, 2003, 12:16:25
Job time : 5.37628 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: July 24, 2003, 11:57:13 ; Search time 5.66764 Seconds
(without alignments)
1026.746 Million cell updates/sec

Title: US-09-991-681-31

Perfect score: 260
Sequence: 1 EPLGPRGQDSPILQRPQHLM.....SAGPELLRQDKRPRSGSTGS 49

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTOS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10C_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	63	24.2	70	9	US-09-864-761-44434
2	56.5	21.7	1000	15	US-10-128-714-3305
3	56	21.5	70	9	US-09-864-761-46074
4	56	21.5	432	15	US-10-156-761-10911
5	55	21.2	32	9	US-09-864-761-35805
6	55	21.2	89	15	US-10-106-698-5596
7	55	21.2	1629	9	US-09-972-467-2
8	55	21.2	1882	10	US-09-918-171A-13
9	55	21.2	1907	10	US-09-938-330-25
10	54.5	21.0	1315	10	US-09-990-046-10
11	54.5	21.0	1433	15	US-10-224-249-14
12	54	20.8	623	14	US-10-108-605-125
13	54	20.8	623	14	US-10-108-605-129
14	54	20.8	638	15	US-10-059-585-8
15	54	20.8	779	8	US-08-817-832B-31

16	54	20.8	793	15	US-10-195-101-32	Sequence 32, Appl
17	54	20.8 <td>795</td> <td>15<td>US-09-919-585-12</td><td>Sequence 12, Appl</td></td>	795	15 <td>US-09-919-585-12</td> <td>Sequence 12, Appl</td>	US-09-919-585-12	Sequence 12, Appl
18	54	20.8 <td>795</td> <td>15<td>US-10-142-156-9</td><td>Sequence 9, Appl</td></td>	795	15 <td>US-10-142-156-9</td> <td>Sequence 9, Appl</td>	US-10-142-156-9	Sequence 9, Appl
19	53.5	20.6	244	10 <td>US-09-734-329-5</td> <td>Sequence 5, Appl</td>	US-09-734-329-5	Sequence 5, Appl
20	53.5	20.6	428	10 <td>US-09-734-329-2</td> <td>Sequence 2, Appl</td>	US-09-734-329-2	Sequence 2, Appl
21	53	20.4	90	9	US-09-867-550-1392	Sequence 1392, Ap
22	53	20.4	323	10 <td>US-09-912-672A-8</td> <td>Sequence 8, Appl</td>	US-09-912-672A-8	Sequence 8, Appl
23	53	20.4	538	10 <td>US-09-976-740-43</td> <td>Sequence 43, Appl</td>	US-09-976-740-43	Sequence 43, Appl
24	53	20.4	538	14 <td>US-10-023-529-43</td> <td>Sequence 43, Appl</td>	US-10-023-529-43	Sequence 43, Appl
25	53	20.4	538	14 <td>US-10-023-523-43</td> <td>Sequence 43, Appl</td>	US-10-023-523-43	Sequence 43, Appl
26	53	20.4	560	10 <td>US-09-912-672A-5</td> <td>Sequence 5, Appl</td>	US-09-912-672A-5	Sequence 5, Appl
27	53	20.4	574	9	US-09-728-911-25	Sequence 25, Appl
28	53	20.4	574	10 <td>US-09-870-574-4</td> <td>Sequence 4, Appl</td>	US-09-870-574-4	Sequence 4, Appl
29	53	20.4	574	10 <td>US-09-912-672A-2</td> <td>Sequence 2, Appl</td>	US-09-912-672A-2	Sequence 2, Appl
30	53	20.4	574	14 <td>US-10-006-867-164</td> <td>Sequence 164, App</td>	US-10-006-867-164	Sequence 164, App
31	53	20.4	574	14 <td>US-10-063-547-164</td> <td>Sequence 164, App</td>	US-10-063-547-164	Sequence 164, App
32	53	20.4	574	15 <td>US-10-063-616-164</td> <td>Sequence 164, App</td>	US-10-063-616-164	Sequence 164, App
33	53	20.4	574	15 <td>US-10-063-502-164</td> <td>Sequence 164, App</td>	US-10-063-502-164	Sequence 164, App
34	53	20.4	574	15 <td>US-10-063-518-164</td> <td>Sequence 164, App</td>	US-10-063-518-164	Sequence 164, App
35	53	20.4	574	15 <td>US-10-063-598-164</td> <td>Sequence 164, App</td>	US-10-063-598-164	Sequence 164, App
36	53	20.4	574	15 <td>US-10-227-693-164</td> <td>Sequence 164, App</td>	US-10-227-693-164	Sequence 164, App
37	53	20.4	574	15 <td>US-10-063-567-164</td> <td>Sequence 164, App</td>	US-10-063-567-164	Sequence 164, App
38	53	20.4	574	15 <td>US-10-063-538-164</td> <td>Sequence 164, App</td>	US-10-063-538-164	Sequence 164, App
39	53	20.4	574	15 <td>US-10-090-365-25</td> <td>Sequence 25, Appl</td>	US-10-090-365-25	Sequence 25, Appl
40	53	20.4	574	15 <td>US-10-233-873A-5</td> <td>Sequence 5, Appl</td>	US-10-233-873A-5	Sequence 5, Appl
41	53	20.4	574	15 <td>US-10-063-599-164</td> <td>Sequence 164, App</td>	US-10-063-599-164	Sequence 164, App
42	53	20.4	574	15 <td>US-10-063-595-164</td> <td>Sequence 164, App</td>	US-10-063-595-164	Sequence 164, App
43	53	20.4	574	15 <td>US-10-104-919-25</td> <td>Sequence 25, Appl</td>	US-10-104-919-25	Sequence 25, Appl
44	53	20.4	574	15 <td>US-10-238-565-4</td> <td>Sequence 4, Appl</td>	US-10-238-565-4	Sequence 4, Appl
45	53	20.4	574	15 <td>US-10-063-580-164</td> <td>Sequence 164, App</td>	US-10-063-580-164	Sequence 164, App

ALIGNMENTS

RESULT 1
US-09-864-761-44434
Sequence 44434, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: A60616-X-1
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US/09/864, 761
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263, 6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30

```

PRIORITY APPLICATION NUMBER: PCT/US01/00662
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00661
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00670
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: US 60/234,687
PRIORITY FILING DATE: 2000-09-21
PRIORITY APPLICATION NUMBER: US 09/608,408
PRIORITY FILING DATE: 2000-06-30
PRIORITY APPLICATION NUMBER: US 09/774,203
PRIORITY FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 44434
LENGTH: 70
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004952.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.73
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.64
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.77
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.85
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.84
US-09-864-761-44434

Query Match          24.2%; Score 63; DB 9; Length 70;
Best Local Similarity 34.5%; Pred. No. 0.78;
Matches 19; Conservative 7; Mismatches 13; Indels 16; Gaps 3;

QY 5 PRGQDSPLQRPQHLMDQGMHRSFSA-----GP-----ELLQDKRRSSSTG 48
DB 12 PPGQ-----QPPHLLHRRGRHRSSEAVSHSRPQGVRIHLQVRODEAGLGATG 61

RESULT 2
US-10-128-714-3305
Sequence 3305, Application US/10128714
GENERAL INFORMATION:
APPLICANT: Hu, Wengf
APPLICANT: Tishkoff, Daniel
APPLICANT: Zamudio, Carlos
APPLICANT: Broshkin, Alexey M
APPLICANT: Lemieux, Sebastien M
TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
FILE REFERENCE: 10182-018-999
CURRENT APPLICATION NUMBER: US/10/128,714
CURRENT FILING DATE: 2002-04-23
PRIORITY APPLICATION NUMBER: US 60/285,697
PRIORITY FILING DATE: 2001-04-23
PRIORITY APPLICATION NUMBER: US 60/287,066
PRIORITY FILING DATE: 2001-04-27
PRIORITY APPLICATION NUMBER: US 60/295,890
PRIORITY FILING DATE: 2001-06-05
PRIORITY APPLICATION NUMBER: US 60/303,899
PRIORITY FILING DATE: 2001-07-09
PRIORITY APPLICATION NUMBER: US 60/316,362
PRIORITY FILING DATE: 2001-08-31
NUMBER OF SEQ ID NOS: 8603
SOFTWARE: Patentin version 3.1
SEQ ID NO 3305
LENGTH: 1000
TYPE: PRT
ORGANISM: Aspergillus fumigatus
US-10-128-714-3305

Query Match          21.7%; Score 56.5; DB 15; Length 1000;
Best Local Similarity 41.7%; Pred. No. 98;

PRIORITY APPLICATION NUMBER: PCT/US01/00662
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00661
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00670
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: US 60/234,687
PRIORITY FILING DATE: 2000-09-21
PRIORITY APPLICATION NUMBER: US 09/608,408
PRIORITY FILING DATE: 2000-06-30
PRIORITY APPLICATION NUMBER: US 09/774,203
PRIORITY FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 46074
LENGTH: 70
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC008752.3
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.7
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.4
OTHER INFORMATION: EST HUMAN HIT: BF314489.1, EVALUATE 2.00e-28
OTHER INFORMATION: SWISSPROT HIT: Q13526, EVALUATE 2.00e-29
US-09-864-761-46074

Matches 15; Conservative 6; Mismatches 12; Indels 3; Gaps 1;

QY 14 QRPQHLMDQGMHRSFSAPELLR---QDKRRSGS 46
DB 713 RKRKRSPPDGQRHRSKSSSEVRSLPRKKRRSGS 748

RESULT 3
US-09-864-761-46074
Sequence 46074, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aeomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIORITY APPLICATION NUMBER: US 60/180,312
PRIORITY FILING DATE: 2000-02-04
PRIORITY APPLICATION NUMBER: US 60/207,456
PRIORITY FILING DATE: 2000-05-26
PRIORITY APPLICATION NUMBER: US 09/632,366
PRIORITY FILING DATE: 2000-08-03
PRIORITY APPLICATION NUMBER: GB 24263.6
PRIORITY FILING DATE: 2000-10-04
PRIORITY APPLICATION NUMBER: US 60/236,359
PRIORITY FILING DATE: 2000-09-27
PRIORITY APPLICATION NUMBER: PCT/US01/00666
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00667
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00664
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00669
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00665
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00668
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00663
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00662
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00661
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00670
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: US 60/234,687
PRIORITY FILING DATE: 2000-09-21
PRIORITY APPLICATION NUMBER: US 09/608,408
PRIORITY FILING DATE: 2000-06-30
PRIORITY APPLICATION NUMBER: US 09/774,203
PRIORITY FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 46074
LENGTH: 70
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC008752.3
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.7
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.4
OTHER INFORMATION: EST HUMAN HIT: BF314489.1, EVALUATE 2.00e-28
OTHER INFORMATION: SWISSPROT HIT: Q13526, EVALUATE 2.00e-29
US-09-864-761-46074
```

Query Match 21.5%; Score 56; DB 9; Length 70;
Best Local Similarity 36.6%; Pred. No. 6;
Matches 15; Conservative 5; Mismatches 17; Indels 4; Gaps 1;

Qy 4 GPRGDSFLORPQHLMDOGMRHSFSAPELLRODKRPRS 44
Db 25 GKNGCGEPARVRCSHLT-----VKHSQSRPSSRWQEKITRT 61

RESULT 4

US-10-156-761-10911
Sequence 10911, Application US/10156761
GENERAL INFORMATION:

APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 10911
LENGTH: 432
TYPE: PRT
ORGANISM: Streptomyces avermitilis
US-10-156-761-10911

Query Match 21.5%; Score 56; DB 15; Length 432;
Best Local Similarity 34.8%; Pred. No. 45;
Matches 16; Conservative 4; Mismatches 10; Indels 16; Gaps 2;

Qy 5 PRGDSFLORPQHLMDOG-----QMRHSFSAPELLROD 39
Db 199 PRROD-----RPDLLETGAVYGMATGFRARHRFFGRTLETVRTD 239

RESULT 5

US-09-864-761-35805
Sequence 35805, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wenheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecm1ca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667

PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 35805
LENGTH: 32
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:

OTHER INFORMATION: MAP TO AC007688.15
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 35
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.8
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 3.9
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.9
OTHER INFORMATION: EST_HUMAN HIT: BE531168.1, EVALU0 3.90e-01
US-09-864-761-35805

Query Match 21.2%; Score 55; DB 9; Length 32;
Best Local Similarity 71.4%; Pred. No. 3.3;
Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 25 MRHSFSAPELLROQ 38
Db 11 LQHSFSAPELLROQ 24

RESULT 6

US-10-106-698-5596
Sequence 5596, Application US/10106698
Publication No. US20030109690A1
GENERAL INFORMATION:

APPLICANT: Ruben et al.
TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
FILE REFERENCE: PA005P1
CURRENT APPLICATION NUMBER: US/10/106,698
CURRENT FILING DATE: 2002-03-27
PRIOR APPLICATION NUMBER: PCT/US00/26224
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US 60/157,137
PRIOR FILING DATE: 1999-09-29
PRIOR APPLICATION NUMBER: US 60/163,280
PRIOR FILING DATE: 1999-11-03
NUMBER OF SEQ ID NOS: 8664
SOFTWARE: PatentIn Ver. 3.0
SEQ ID NO 5596
LENGTH: 89

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-106-698-5596

Query Match
Best Local Similarity 21.2%; Score 55; DB 15; Length 89;
Matches 17; Conservative 8; Mismatches 17; Indels 6; Gaps 3;

QY 2 PLPGQGDSPLLQRPQHL-MDQGQMRH--SFSGAPELLRODKRPRSGS 46
DB 9 PLSPWKKRTNVEPEGVQLDQGDILHLTVFSVCPSLY---SNVRNGS 53

RESULT 7
US-09-972-467-2
; Sequence 2, Application US/09972467
; Patent No. US20020090373A1
; GENERAL INFORMATION:
; APPLICANT: PRIZER INC.
; TITLE OF INVENTION: ADAMTS POLYPEPTIDES, NUCLEIC ACIDS ENCODING THEM, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: PC10850A
; CURRENT APPLICATION NUMBER: US/09/972,467
; CURRENT FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1629
; TYPE: PRT
; ORGANISM: Human
US-09-972-467-2

Query Match
Best Local Similarity 21.2%; Score 55; DB 9; Length 1629;
Matches 14; Conservative 3; Mismatches 16; Indels 6; Gaps 1;

QY 8 QDSPLLQRPQHLMDOGMKHSFSAGPELLRODKRPRSGS 46
DB 1288 QDCSMSPCPORTPDSGLAQHPFO-----NEDYRPRSAS 1320

RESULT 8
US-09-918-171A-13
; Sequence 13, Application US/09918171A
; Patent No. US20020110894A1
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirahara, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/0419
; CURRENT APPLICATION NUMBER: US/09/918,171A
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 09/369,364
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 1882
; TYPE: PRT
; ORGANISM: Homo sapiens ADAMTS-9
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (468)
; OTHER INFORMATION: Xaa = Cys
; NAME/KEY: MOD RES
; LOCATION: (521)
; OTHER INFORMATION: Xaa = Tyr
US-09-918-171A-13

Query Match
Best Local Similarity 21.2%; Score 55; DB 10; Length 1882;
Matches 14; Conservative 3; Mismatches 16; Indels 6; Gaps 1;
```

```

QY 8 QDSPLLQRPQHLMDOGMKHSFSAGPELLRODKRPRSGS 46
DB 1235 QDCSMSPCPORTPDSGLAQHPFO-----NEDYRPRSAS 1267

RESULT 9
US-09-938-330-25
; Sequence 25, Application US/09938330
; Patent No. US20020115838A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hilburn, Erin
; APPLICANT: Scoville, John
; APPLICANT: Fridde, Carl Johan
; APPLICANT: Hu, Yi
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020115838A1el Human Proteases and Polynucleotides Encodin
; FILE REFERENCE: LEX-0237-USA
; CURRENT APPLICATION NUMBER: US/09/938,330
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: US 60/227,104
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60/233,796
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 1907
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-938-330-25

Query Match
Best Local Similarity 21.2%; Score 55; DB 10; Length 1907;
Matches 14; Conservative 3; Mismatches 16; Indels 6; Gaps 1;
```

```

QY 8 QDSPLLQRPQHLMDOGMKHSFSAGPELLRODKRPRSGS 46
DB 1260 QDCSMSPCPORTPDSGLAQHPFO-----NEDYRPRSAS 1292

RESULT 10
US-09-990-046-10
; Sequence 10, Application US/09990046
; Patent No. US20020156245A1
; GENERAL INFORMATION:
; APPLICANT: de Sauvage, Frederic
; APPLICANT: Carpenter, David A.
; TITLE OF INVENTION: Patched-2
; FILE REFERENCE: P1405R1
; CURRENT APPLICATION NUMBER: US/09/990,046
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/293,505
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-15
; NUMBER OF SEQ ID NOS: 32
; SEQ ID NO 10
; LENGTH: 1315
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-990-046-10

Query Match
Best Local Similarity 21.0%; Score 54.5; DB 10; Length 1315;
Matches 15; Conservative 11; Mismatches 19; Indels 11; Gaps 2;

QY 3 LGPRGQDSPLL-QRPQHLMDOGMKHSFSAGPELLRODKR-----PRSGST 47
DB 292 LAPKGNOSRILTLQAYKMAEAMQKQKONTGPALBQEDKTSKVAIPGTAPRLRGAT 347

RESULT 11
US-10-224-249-14
```



```
Sequence 14, Application US/10224249
; Publication No. US20030087867A1
; GENERAL INFORMATION:
; APPLICANT: Vogels, Ronald V.
; APPLICANT: Verlinden, Stefan F. P.
; TITLE OF INVENTION: Gene therapy for enhancing and/or inducing angiogenesis
; FILE REFERENCE: 2183-5233US
; CURRENT APPLICATION NUMBER: US/10/224,249
; CURRENT FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: PCT/NL00/00482
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: EP 99202263.2
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: US 60/143,101
; PRIOR FILING DATE: 1999-07-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 1433
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; LOCATION: (1)..(1433)
; OTHER INFORMATION: Human nitric oxide synthase
US-10-224-249-14
```

```
Query Match          21.0%; Score 54.5; DB 15; Length 1433;
Best Local Similarity 29.0%; Pred. No. 2.6e+02;
Matches 18; Conservative 5; Mismatches 16; Indels 23; Gaps 3;
```

```
Qy 1 EPLG-----PRQDSPL-----QRPOLMDQGMHSE---SAGPELLR 37
Db 124 QELGPPTKAVDLSDHPKKEQPLAVDASGPGNGPQAHYDGGQASGLPHANGWQAPR 183
```

```
Qy 38 QD 39
Db 184 QD 185
```

```
RESULT 12
US-10-108-605-125
; Sequence 125, Application US/10108605
; Publication No. US20020160934A1
; GENERAL INFORMATION:
; APPLICANT: Broadus, Julie
; APPLICANT: Stam, Lynn
; APPLICANT: Bachmann, Jane
; APPLICANT: Kamdar, Kim
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCODE
; TITLE OF INVENTION: PROTEINS ESSENTIAL FOR LARVAL VIABILITY AND USES THEREOF
; FILE REFERENCE: 31133B
; CURRENT APPLICATION NUMBER: US/10/108,605
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US 09/761,142
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/176,418
; PRIOR FILING DATE: 2000-01-14
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 125
; LENGTH: 623
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-108-605-125
```

```
Query Match          20.8%; Score 54; DB 14; Length 623;
Best Local Similarity 35.4%; Pred. No. 1.2e+02;
Matches 17; Conservative 4; Mismatches 19; Indels 8; Gaps 2;
```

```
Qy 8 QDSPLLRPPQHL-----MDQGMHSPSAGPELLRQDKRPRSSGTGS 49
Db 110 QQSLLLRHPQQQQQSHSQSQQQQQHGYSQAQL--PHHRLSGSGSTGS 155
```

```
RESULT 13
US-10-108-605-129
; Sequence 129, Application US/10108605
; Publication No. US20020160934A1
; GENERAL INFORMATION:
; APPLICANT: Broadus, Julie
; APPLICANT: Stam, Lynn
; APPLICANT: Bachmann, Jane
; APPLICANT: Kamdar, Kim
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCODE
; TITLE OF INVENTION: PROTEINS ESSENTIAL FOR LARVAL VIABILITY AND USES THEREOF
; FILE REFERENCE: 31133B
; CURRENT APPLICATION NUMBER: US/10/108,605
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US 09/761,142
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/176,418
; PRIOR FILING DATE: 2000-01-14
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 129
; LENGTH: 623
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-108-605-129
```

```
Query Match          20.8%; Score 54; DB 14; Length 623;
Best Local Similarity 35.4%; Pred. No. 1.2e+02;
Matches 17; Conservative 4; Mismatches 19; Indels 8; Gaps 2;
```

```
Qy 8 QDSPLLRPPQHL-----MDQGMHSPSAGPELLRQDKRPRSSGTGS 49
Db 110 QQSLLLRHPQQQQQSHSQSQQQQQHGYSQAQL--PHHRLSGSGSTGS 155
```

```
RESULT 14
US-10-059-585-8
; Sequence 8, Application US/10059585
; Publication No. US20030082776A1
; GENERAL INFORMATION:
; APPLICANT: Ota, Toshio
; APPLICANT: Isogai, Takao
; APPLICANT: Nishikawa, Tetsuo
; APPLICANT: Hayashi, Koji
; APPLICANT: Otsuka, Kaoru
; APPLICANT: Yamamoto, Jun-ichi
; APPLICANT: Ishii, Shizuko
; APPLICANT: Sugiyama, Tomoyasu
; APPLICANT: Wakamatsu, Ai
; APPLICANT: Nagai, Keiichi
; APPLICANT: Otsuki, Tetsuji
; APPLICANT: Funahashi, Shin-Ichi
; APPLICANT: Senoo, Chiaki
; APPLICANT: Nezu, Jun-ichi
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN
; TITLE OF INVENTION: KINASE/PROTEIN PHOSPHATASE
; FILE REFERENCE: 06501-098001
; CURRENT APPLICATION NUMBER: US/10/059,585
; CURRENT FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: PCT/JP00/05060
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/183,322
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: US 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 11-248036
; PRIOR FILING DATE: 1999-07-29
```

NUMBER OF SEQ ID NOS: 64
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO: 8
 LENGTH: 638
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-059-585-8

Query Match 20.8%; Score 54; DB 15; Length 638;
 Best Local Similarity 36.2%; Pred. No. 1.2e+02;
 Matches 17; Conservative 5; Mismatches 15; Indels 10; Gaps 2;

QY 13 LORPQHL-----MDQGMRHSFSAGPEL---LRQDKRRSGSGTGS 49
 DB 266 LQSPALHKYQRISANQKQRRFSDHAGPSIPPAVSYTKRPQANVSVE 312

RESULT 15

US-08-817-832B-31
 ; Sequence 31, Application US/08817832B
 ; Publication No. US20030104516A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MANDELKOW, Eckhard, et al.
 ; TITLE OF INVENTION: No. US20030104516A1 Protein Kinase (NPK-110)
 ; NUMBER OF SEQUENCES: 32
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 ; STREET: 233 S. Wacker Drive, 6300 Sears Tower
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: US
 ; ZIP: 60606-6402
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/817, 832B
 ; FILING DATE: 28-APR-1997
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: WO PCT/EP95/04258
 ; FILING DATE: 30-OCT-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: EP 94 11 7122.5
 ; FILING DATE: 28-OCT-1994
 ; INFORMATION FOR SEQ ID NO: 31:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 779 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-817-832B-31

Query Match 20.8%; Score 54; DB 8; Length 779;
 Best Local Similarity 36.2%; Pred. No. 1.5e+02;
 Matches 17; Conservative 5; Mismatches 15; Indels 10; Gaps 2;

QY 13 LORPQHL-----MDQGMRHSFSAGPEL---LRQDKRRSGSGTGS 49
 DB 388 LQSPALHKYQRISANQKQRRFSDHAGPSIPPAVSYTKRPQANVSVE 434

Search completed: July 24, 2003, 12:18:21
 Job time : 6.86764 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 11:49:53 ; Search time 73.8228 Seconds
(without alignments)
577.701 Million cell updates/sec

Title: US-09-991-681-31

Perfect score: 260
Sequence: 1 EPLGPRGQDSPILQRPQHLM.....SAGPELLRQDKRPRSGSTGS 49

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 5580241 seqs, 870357830 residues

Total number of hits satisfying chosen parameters: 5580241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending Patents AA Main:*

1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097A_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US097B_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US099A_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US099B_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
28: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
29: /cgn2_6/ptodata/1/paa/US103_COMB.pep.*
30: /cgn2_6/ptodata/1/paa/US104_COMB.pep.*
31: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	260	100.0	49	US-08-842-385-10	Sequence 10, Appl
2	260	100.0	49	US-09-991-681-31	Sequence 31, Appl

3	260	100.0	467	12	US-08-842-385-6	Sequence 6, Appl
4	260	100.0	518	25	US-09-991-681-27	Sequence 27, Appl
5	260	100.0	1770	1	PCT-US03-01943-44	Sequence 44, Appl
6	260	100.0	1770	27	US-10-144-198-44	Sequence 44, Appl
7	260	100.0	1807	1	PCT-US01-08631-40087	Sequence 40087, A
8	260	100.0	1839	1	PCT-US01-42950-495	Sequence 495, App
9	260	100.0	1839	30	US-10-416-993-495	Sequence 495, App
10	260	100.0	1872	1	PCT-US03-04508-32	Sequence 32, Appl
11	260	100.0	1982	1	PCT-US03-04508-32	Sequence 40090, A
12	260	100.0	2221	1	PCT-US03-01943-30	Sequence 30, Appl
13	260	100.0	2221	27	US-10-144-198-30	Sequence 30, Appl
14	70.5	27.1	309	30	US-10-419-128-17975	Sequence 17975, A
15	68	26.2	133	30	US-10-437-963-135832	Sequence 135832, A
16	67	25.8	195	1	PCT-US01-08656-9053	Sequence 9053, Ap
17	65	25.0	693	30	US-10-419-128-24059	Sequence 24059, A
18	64.5	24.8	501	30	US-10-419-128-21596	Sequence 21596, A
19	63.5	24.4	132	22	US-09-758-446-1669	Sequence 1669, Ap
20	63.5	24.4	132	28	US-10-212-759-1669	Sequence 1669, Ap
21	63	24.2	70	1	PCT-US01-00663-34316	Sequence 34316, A
22	63	24.2	70	23	US-09-864-761-44434	Sequence 44434, A
23	63	24.2	70	27	US-10-182-993-33256	Sequence 33256, A
24	63	24.2	70	28	US-10-203-134-34171	Sequence 34171, A
25	63	24.2	70	28	US-10-203-135-33418	Sequence 33418, A
26	63	24.2	70	28	US-10-203-136-34264	Sequence 34264, A
27	63	24.2	70	28	US-10-203-137-34316	Sequence 34316, A
28	63	24.2	70	28	US-10-203-139-32981	Sequence 32981, A
29	63	24.2	171	1	PCT-US01-00631-52610	Sequence 52610, A
30	63	24.2	504	30	US-10-419-128-23374	Sequence 23374, A
31	62.5	24.0	210	30	US-10-424-599-180951	Sequence 180951, A
32	62	23.8	87	20	US-09-644-265-116	Sequence 116, App
33	62	23.8	87	28	US-10-263-828-116	Sequence 116, App
34	62	23.8	567	30	US-10-419-128-28435	Sequence 28435, A
35	61.5	23.7	407	31	US-60-452-680-14462	Sequence 14462, A
36	61.5	23.7	407	31	US-60-453-050-9016	Sequence 9016, Ap
37	61.5	23.7	407	31	US-60-453-135-9016	Sequence 9016, Ap
38	61.5	23.7	407	31	US-60-466-412-9016	Sequence 9016, Ap
39	61.5	23.7	833	30	US-10-437-963-183706	Sequence 183706, A
40	61	23.5	171	30	US-10-419-128-18380	Sequence 18380, A
41	61	23.5	179	30	US-10-437-963-179732	Sequence 179732, A
42	61	23.5	250	1	PCT-US02-30474-1778	Sequence 1778, Ap
43	61	23.5	250	26	US-10-097-105-1562	Sequence 1562, Ap
44	61	23.5	250	31	US-60-324-631-1783	Sequence 1783, Ap
45	61	23.5	250	31	US-60-453-050-13346	Sequence 13346, A

ALIGNMENTS

RESULT 1
US-08-842-385-10
Sequence 10 Application US/08842385

GENERAL INFORMATION:
APPLICANT: Russell, John
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL FOR DETECTING DISEASE OF THE PROSTATE
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESS: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/842,385
FILING DATE:
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Potembski, Priscilla E
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 6084.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/937-6365
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: None
US-08-842-385-10

Query Match 100.0%; Score 260; DB 12; Length 49;
Best Local Similarity 100.0%; Pred. No. 6.6e-25;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EPLGPGDPSPLQRPQHLMQGMHRSFSAAGPELLRQDKRPRSSTGS 49
Db 1 EPLGPGDPSPLQRPQHLMQGMHRSFSAAGPELLRQDKRPRSSTGS 49

RESULT 2
US-09-991-681-31
Sequence 31, Application US/09991681
GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA
COHEN, MAURICE
COLPITTS, TRACEY L.
FRIEDMAN, PAULA N.
GORDON, JULIAN
GRANADOS, EDWARD N.
HODGES, STEVEN C.
KLASS, MICHAEL R.
KRATOCHVIL, JON D.
ROBERTS-RAP, LISA
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSER: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/991,681
FILING DATE: 26-Nov-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/065,383
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6084.US.PI
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: None
SEQUENCE DESCRIPTION: SEQ ID NO: 31:
US-09-991-681-31

Query Match 100.0%; Score 260; DB 25; Length 49;
Best Local Similarity 100.0%; Pred. No. 6.6e-25;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EPLGPGDPSPLQRPQHLMQGMHRSFSAAGPELLRQDKRPRSSTGS 49
Db 1 EPLGPGDPSPLQRPQHLMQGMHRSFSAAGPELLRQDKRPRSSTGS 49

RESULT 3
US-08-842-385-6
Sequence 6, Application US/08842385
GENERAL INFORMATION:
APPLICANT: Russell, John
COLPITTS, TRACEY
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
FOR DETECTING DISEASE OF THE PROSTATE
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSER: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/842,385
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Potembski, Priscilla E
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 6084.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/937-6365
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: None
US-08-842-385-6

Query Match 100.0%; Score 260; DB 12; Length 467;
Best Local Similarity 100.0%; Pred. No. 1.1e-23;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EPLGPGDPSPLQRPQHLMQGMHRSFSAAGPELLRQDKRPRSSTGS 49
Db 342 EPLGPGDPSPLQRPQHLMQGMHRSFSAAGPELLRQDKRPRSSTGS 390

```
RESULT 4
US-09-991-681-27
; Sequence 27, Application US/09991681
; GENERAL INFORMATION:
; APPLICANT: BILLING-MEDLI, PATRICIA
; COHEN, MAURICE
; COLPITTS, TRACEY L.
; FRIEDMAN, PAULA N.
; GORDON, JULIAN
; GRANADOS, EDWARD N.
; HODGES, STEVEN C.
; KLAAS, MICHAEL R.
; KRATOCHVIL, JON D.
; ROBERTS-RAPP, LISA
; TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
; FOR DETECTING DISEASES OF THE PROSTATE
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: PASTESEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/991,681
; FILING DATE: 26-Nov-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/065,383
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Becker, Cheryl L.
; REGISTRATION NUMBER: 35,441
; REFERENCE/DOCKET NUMBER: 6084.US.P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847/935-1729
; TELEFAX: 847/938-2623
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 518 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: None
; SEQUENCE DESCRIPTION: SEQ ID NO: 27:
US-09-991-681-27

Query Match 100.0%; Score 260; DB 25; Length 518;
Best Local Similarity 100.0%; Pred. No. 1.2e-23;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 EPLGPGDPSPLORPOHLMDOGMRHSFSAGPELLRQDKRPRSGSTGS 49
Db 393 EPLGPGDPSPLORPOHLMDOGMRHSFSAGPELLRQDKRPRSGSTGS 441

RESULT 5
PCT-US03-01943-44
; Sequence 44, Application PC/TUS0301943
; GENERAL INFORMATION:
; APPLICANT: ORIGENE TECHNOLOGIES INC
; TITLE OF INVENTION: CANCER GENES
; FILE REFERENCE: 3U 9U 901 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/01943
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: US 10/054,935
```

```
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: US 60/356,130
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 10/102,946
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: US 10/117,229
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: US 10/144,198
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: US 10/197,824
; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44
; LENGTH: 1770
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-01943-44

Query Match 100.0%; Score 260; DB 1; Length 1770;
Best Local Similarity 100.0%; Pred. No. 5.8e-23;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 EPLGPGDPSPLORPOHLMDOGMRHSFSAGPELLRQDKRPRSGSTGS 49
Db 1645 EPLGPGDPSPLORPOHLMDOGMRHSFSAGPELLRQDKRPRSGSTGS 1693

RESULT 6
US-10-144-198-44
; Sequence 44, Application US/10144198
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies Inc
; TITLE OF INVENTION: Regulated Prostate Cance Genes
; FILE REFERENCE: 9U 105 R1
; CURRENT APPLICATION NUMBER: US/10/144,198
; CURRENT FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 44
; LENGTH: 1770
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-198-44

Query Match 100.0%; Score 260; DB 27; Length 1770;
Best Local Similarity 100.0%; Pred. No. 5.8e-23;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 EPLGPGDPSPLORPOHLMDOGMRHSFSAGPELLRQDKRPRSGSTGS 49
Db 1645 EPLGPGDPSPLORPOHLMDOGMRHSFSAGPELLRQDKRPRSGSTGS 1693

RESULT 7
PCT-US01-08631-40087
; Sequence 40087, Application PC/TUS0108631
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-049
; CURRENT APPLICATION NUMBER: PCT/US01/08631
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 40087
; LENGTH: 1807
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (48)..(62)
OTHER INFORMATION: Fusion glycoprotein F0 domain identified by eMATRIX,
OTHER INFORMATION: accession number PF00523D, p-value=7.188e-10, raw score of 11.39
NAME/KEY: DOMAIN
LOCATION: (941)..(950)
OTHER INFORMATION: Helper component proteinase domain identified by Pfam,
OTHER INFORMATION: accession name Peptidase_C6, E-value=0.0056, Pfam score of 7.9
PCT-US01-08631-40087
```

```
Query Match      100.0%; Score 260; DB 1; Length 1807;
Best Local Similarity 100.0%; Pred. No. 6e-23;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 EPLGPGGDSPLLRPOHLMDOGQMRHSFSAGPELLRQDKRPRSGSTGS 49
Db      1682 EPLGPGGDSPLLRPOHLMDOGQMRHSFSAGPELLRQDKRPRSGSTGS 1730
```

```
RESULT 8
PCT-US01-42950-495
Sequence 495, Application PC/TUS0142950
```

```
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 21272-096
```

```
CURRENT APPLICATION NUMBER: PCT/US01/42950
```

```
CURRENT FILING DATE: 2001-11-16
```

```
PRIOR APPLICATION NUMBER: 09/714,936
```

```
PRIOR FILING DATE: 2000-11-17
```

```
NUMBER OF SEQ ID NOS: 682
```

```
SOFTWARE: Patentin version 3.0
```

```
SEQ ID NO 495
```

```
LENGTH: 1839
```

```
TYPE: PRT
```

```
ORGANISM: Homo sapiens
```

```
PCT-US01-42950-495
```

```
Query Match      100.0%; Score 260; DB 1; Length 1839;
Best Local Similarity 100.0%; Pred. No. 6.1e-23;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 EPLGPGGDSPLLRPOHLMDOGQMRHSFSAGPELLRQDKRPRSGSTGS 49
Db      1714 EPLGPGGDSPLLRPOHLMDOGQMRHSFSAGPELLRQDKRPRSGSTGS 1762
```

```
RESULT 9
US-10-416-993-495
Sequence 495, Application US/10416993
```

```
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
```

```
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
```

```
FILE REFERENCE: 21272-096
```

```
CURRENT APPLICATION NUMBER: US/10/416,993
```

```
CURRENT FILING DATE: 2003-11-16
```

```
PRIOR APPLICATION NUMBER: 09/714,936
```

```
PRIOR FILING DATE: 2000-11-17
```

```
NUMBER OF SEQ ID NOS: 682
```

```
SOFTWARE: Patentin version 3.0
```

```
SEQ ID NO 495
```

```
LENGTH: 1839
```

```
TYPE: PRT
```

```
ORGANISM: Homo sapiens
```

```
PCT-US-10-416-993-495
```

```
Query Match      100.0%; Score 260; DB 30; Length 1839;
Best Local Similarity 100.0%; Pred. No. 6.1e-23;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 EPLGPGGDSPLLRPOHLMDOGQMRHSFSAGPELLRQDKRPRSGSTGS 49
Db      1 EPLGPGGDSPLLRPOHLMDOGQMRHSFSAGPELLRQDKRPRSGSTGS 49
```

```
Db      1714 EPLGPGGDSPLLRPOHLMDOGQMRHSFSAGPELLRQDKRPRSGSTGS 1762
```

```
RESULT 10
PCT-US03-04508-32
Sequence 32, Application PC/TUS0304508
```

```
GENERAL INFORMATION:
APPLICANT: IDEC PHARMACEUTICALS
```

```
APPLICANT: GATELY, DENNIS
```

```
TITLE OF INVENTION: PROSTATE SPECIFIC GENES AND THE USE THEREOF IN DESIGN
```

```
FILE REFERENCE: 037003/0301985
```

```
CURRENT APPLICATION NUMBER: PCT/US03/04508
```

```
CURRENT FILING DATE: 2003-02-19
```

```
PRIOR APPLICATION NUMBER: 60/357,140
```

```
PRIOR FILING DATE: 2002-02-19
```

```
PRIOR APPLICATION NUMBER: 60/396,082
```

```
PRIOR FILING DATE: 2002-07-17
```

```
PRIOR APPLICATION NUMBER: 60/386,759
```

```
PRIOR FILING DATE: 2002-06-10
```

```
NUMBER OF SEQ ID NOS: 93
```

```
SOFTWARE: Patentin Ver. 2.1
```

```
SEQ ID NO 32
```

```
LENGTH: 1872
```

```
TYPE: PRT
```

```
ORGANISM: Homo sapiens
```

```
PCT-US03-04508-32
```

```
Query Match      100.0%; Score 260; DB 1; Length 1872;
Best Local Similarity 100.0%; Pred. No. 6.2e-23;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 EPLGPGGDSPLLRPOHLMDOGQMRHSFSAGPELLRQDKRPRSGSTGS 49
Db      1747 EPLGPGGDSPLLRPOHLMDOGQMRHSFSAGPELLRQDKRPRSGSTGS 1795
```

```
RESULT 11
PCT-US01-08631-40090
```

```
Sequence 40090, Application PC/TUS0108631
```

```
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
```

```
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
```

```
FILE REFERENCE: 21272-049
```

```
CURRENT APPLICATION NUMBER: PCT/US01/08631
```

```
CURRENT FILING DATE: 2001-03-30
```

```
PRIOR APPLICATION NUMBER: 09/540,217
```

```
PRIOR FILING DATE: 2000-03-31
```

```
PRIOR APPLICATION NUMBER: 09/649,167
```

```
PRIOR FILING DATE: 2000-08-23
```

```
NUMBER OF SEQ ID NOS: 60736
```

```
SOFTWARE: Custom
```

```
SEQ ID NO 40090
```

```
LENGTH: 1982
```

```
TYPE: PRT
```

```
ORGANISM: Homo sapiens
```

```
FEATURE:
NAME/KEY: DOMAIN
```

```
LOCATION: (11)..(25)
```

```
OTHER INFORMATION: Fusion glycoprotein F0 domain identified by eMATRIX,
```

```
OTHER INFORMATION: accession number PF00523D, p-value=7.188e-10, raw score of 11.39
```

```
NAME/KEY: DOMAIN
```

```
LOCATION: (1065)..(1074)
```

```
OTHER INFORMATION: Helper component proteinase domain identified by Pfam,
```

```
OTHER INFORMATION: accession name Peptidase_C6, E-value=0.0056, Pfam score of 7.9
```

```
PCT-US01-08631-40090
```

```
Query Match      100.0%; Score 260; DB 1; Length 1982;
Best Local Similarity 100.0%; Pred. No. 6.7e-23;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 EPLGPGGDSPLLRPOHLMDOGQMRHSFSAGPELLRQDKRPRSGSTGS 49
Db      1 EPLGPGGDSPLLRPOHLMDOGQMRHSFSAGPELLRQDKRPRSGSTGS 49
```

Db 1796 EPLGRGDSPLLRPOHLMDOGQNRHSFSAGPELLRODKRPRSGTGS 1844

RESULT 12

PCT-US03-01943-30
; Sequence 30, Application PC/TUS0301943
; GENERAL INFORMATION:
; APPLICANT: ORIGENE TECHNOLOGIES INC
; TITLE OF INVENTION: CANCER GENES
; FILE REFERENCE: 3U 9U 901 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/01943
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: US 10/054,935
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: US 60/356,130
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 10/102,946
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: US 10/117,229
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: US 10/144,198
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: US 10/197,824
; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30
; LENGTH: 2221
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-01943-30

Query Match 100.0%; Score 260; DB 1; Length 2221;
Best Local Similarity 100.0%; Pred. No. 7.7e-23;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EPLGRGDSPLLRPOHLMDOGQNRHSFSAGPELLRODKRPRSGTGS 49
Db 2096 EPLGRGDSPLLRPOHLMDOGQNRHSFSAGPELLRODKRPRSGTGS 2144

RESULT 13

US-10-144-198-30
; Sequence 30, Application US/10144198
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies Inc
; TITLE OF INVENTION: Regulated Prostate Cance Genes
; FILE REFERENCE: 9U 105 R1
; CURRENT APPLICATION NUMBER: US/10/144,198
; CURRENT FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 30
; LENGTH: 2221
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-198-30

Query Match 100.0%; Score 260; DB 27; Length 2221;
Best Local Similarity 100.0%; Pred. No. 7.7e-23;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EPLGRGDSPLLRPOHLMDOGQNRHSFSAGPELLRODKRPRSGTGS 49
Db 2096 EPLGRGDSPLLRPOHLMDOGQNRHSFSAGPELLRODKRPRSGTGS 2144

RESULT 14

US-10-419-128-17975
; Sequence 17975, Application US/10419128
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/10/419,128
; CURRENT FILING DATE: 2003-04-21/252,991
; PRIOR APPLICATION NUMBER: US/09/252,991
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17975
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-419-128-17975

Query Match 27.1%; Score 70.5; DB 30; Length 309;
Best Local Similarity 41.5%; Pred. No. 6.4;
Matches 22; Conservative 3; Mismatches 19; Indels 9; Gaps 2;

Qy 2 PLGRGDSPLLRPOHLMDOGQNRHSFSAGPELLRODKRPRSG 45
Db 194 PLGRGDSPLLRPOHLMDOGQNRHSFSAGPELLRODKRPRSG 246

RESULT 15

US-10-437-963-135832
; Sequence 135832, Application US/10437963
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 135832
; LENGTH: 133
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_3746C.1.pep
US-10-437-963-135832

Query Match 26.2%; Score 68; DB 30; Length 133;
Best Local Similarity 27.8%; Pred. No. 4.6;
Matches 25; Conservative 6; Mismatches 17; Indels 42; Gaps 3;

Qy 2 PLGRGDSPLLRPOHLMDOGQNRHSFSAGPELLRODKRPRSG 45
Db 40 PLGRGDSPLLRPOHLMDOGQNRHSFSAGPELLRODKRPRSG 99
Qy 22 QGMRHSFSAGPELLRODKRPRSGTGS 49
Db 100 KGPARNLSLPSAQTSLDEHRLPLSGATES 129

Search completed: July 24, 2003, 12:15:15
Job time : 74.8228 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: July 24, 2003, 11:54:33 ; Search time 1.43485 Seconds
(without alignments)
140.227 Million cell updates/sec

Title: US-09-991-681-31

Perfect score: 260
Sequence: 1 EPLGPRGDSPLRPOHLM.....SAGPELLRQDKRPRSGSTGS 49

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 41799 seqs, 4106219 residues

Total number of hits satisfying chosen parameters: 41799

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New.*
1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	260	100.0	1872	6	US-10-367-978-32 Sequence 32, App1
2	67	25.8	195	6	US-10-273-573-9053 Sequence 9053, Ap
3	56	21.5	149	6	US-10-273-573-8476 Sequence 8476, Ap
4	56	21.5	205	6	US-10-273-573-9052 Sequence 9052, Ap
5	52	20.0	1239	5	US-09-291-417D-13 Sequence 13, App1
6	51	19.6	139	6	US-10-273-573-7947 Sequence 7947, Ap
7	50.5	19.4	167	6	US-10-273-573-8693 Sequence 8693, Ap
8	50.5	19.4	1233	5	US-09-291-417D-89 Sequence 89, App1
9	50	19.2	501	6	US-10-273-573-7079 Sequence 7079, Ap
10	50	19.2	501	6	US-10-273-573-7748 Sequence 7748, Ap
11	50	19.2	603	1	PCT-US02-18638A-186 Sequence 186, App
12	50	19.2	87	6	US-10-273-573-8772 Sequence 8772, Ap
13	49	18.8	205	6	US-10-464-368-49 Sequence 49, App1
14	49	18.8	1888	6	US-10-273-573-7597 Sequence 7597, Ap
15	49	18.8	1888	6	US-10-273-573-7597 Sequence 7597, Ap
16	49	18.8	1888	6	US-10-273-573-7597 Sequence 7597, Ap
17	48	18.5	525	6	US-10-294-433-740 Sequence 740, App
18	48	18.5	1026	6	US-10-294-433-793 Sequence 793, App
19	48	18.5	1121	7	US-60-478-196-3008 Sequence 3008, App
20	47	18.1	96	6	US-10-273-573-10953 Sequence 10953, A
21	47	18.1	393	6	US-10-391-363A-7 Sequence 7, App1
22	47	18.1	2153	6	US-10-273-573-10697 Sequence 10697, A
23	47	18.1	6820	1	PCT-US03-18787-101 Sequence 101, App
24	47	18.1	6820	1	PCT-US03-18787-101 Sequence 101, App
25	46	17.7	207	6	US-10-273-573-7584 Sequence 7584, Ap
26	46	17.7	827	7	US-60-478-196-3004 Sequence 3004, Ap

27	46	17.7	6842	1	PCT-US03-18787-131 Sequence 131, App
28	46	17.7	6842	1	PCT-US03-19069-131 Sequence 131, App
29	45.5	17.5	319	1	PCT-US03-19589-12 Sequence 12, App1
30	45.5	17.5	390	6	US-10-273-573-9127 Sequence 9127, Ap
31	45.5	17.5	513	6	US-10-273-573-6751 Sequence 6751, Ap
32	45.5	17.5	704	7	US-60-478-196-3006 Sequence 3006, Ap
33	45.5	17.5	718	1	PCT-US03-19660-26 Sequence 26, App1
34	45	17.3	124	6	US-10-273-573-6194 Sequence 6194, Ap
35	45	17.3	185	6	US-10-273-573-10415 Sequence 10415, A
36	45	17.3	379	6	US-10-273-573-6457 Sequence 6457, Ap
37	45	17.3	591	5	US-09-291-417D-103 Sequence 103, App
38	45	17.3	764	6	US-10-294-433-328 Sequence 328, App
39	45	17.3	1550	6	US-10-273-573-6391 Sequence 6391, Ap
40	45	17.3	4074	6	US-10-358-063-1 Sequence 1, App1
41	44.5	17.1	723	7	US-60-479-073-457 Sequence 457, App
42	44.5	17.1	1994	6	US-10-294-433-338 Sequence 338, App
43	44.5	17.1	2041	6	US-10-273-573-8282 Sequence 8282, Ap
44	44	16.9	141	6	US-10-273-573-8834 Sequence 8834, Ap
45	44	16.9	239	6	US-10-273-573-8834 Sequence 8834, Ap

ALIGNMENTS

```
RESULT 1
US-10-367-978-32
Sequence 32, Application US/10367978
GENERAL INFORMATION:
APPLICANT: GATELY, DENNIS
TITLE OF INVENTION: PROSTATE SPECIFIC GENES AND THE USE THEREOF IN DESIGN
FILE REFERENCE: 037003-0301988
CURRENT APPLICATION NUMBER: US/10/367,978
PRIOR FILING DATE: 2003-02-19
PRIOR APPLICATION NUMBER: 60/357,140
PRIOR FILING DATE: 2002-02-19
PRIOR APPLICATION NUMBER: 60/396,082
PRIOR FILING DATE: 2002-07-17
PRIOR APPLICATION NUMBER: 60/386,759
PRIOR FILING DATE: 2002-06-10
NUMBER OF SEQ ID NOS: 89
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 32
LENGTH: 1872
TYPE: PRT
ORGANISM: Homo sapiens
US-10-367-978-32
Query Match
Best Local Similarity 100.0%; Score 260; DB 6; Length 1872;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 EPLGPRGDSPLRPOHLMDOGMHSFSAAGPELLRQDKRPRSGSTGS 49
Db 1747 EPLGPRGDSPLRPOHLMDOGMHSFSAAGPELLRQDKRPRSGSTGS 1795
RESULT 2
US-10-273-573-9053
Sequence 9053, Application US/10273573
GENERAL INFORMATION:
APPLICANT: Hyaseq, Inc
TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 21272-066
CURRENT APPLICATION NUMBER: US/10/273,573
CURRENT FILING DATE: 2002-10-18
PRIOR APPLICATION NUMBER: 09/522,929
PRIOR FILING DATE: 2000-04-18
PRIOR APPLICATION NUMBER: 09/770,160
PRIOR FILING DATE: 2001-01-26
NUMBER OF SEQ ID NOS: 10994
SOFTWARE: Custom
SEQ ID NO 9053
```

```
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (33)..(48)
; OTHER INFORMATION: MW/rsp5/wmp domain proteins domain identified by eMATRIX.
; OTHER INFORMATION: accession number B01159, p-value=3.077e-15, raw score of 13.85
US-10-273-573-9053

Query Match
Best Local Similarity 37.5%; Score 67; DB 6; Length 195;
Matches 18; Conservative 6; Mismatches 18; Indels 6; Gaps 2;

Qy 4 GPGGDSPLQRPQHLMDQGMRHSFSAPELLRQDK--RPRSGSTGS 49
Db 55 GKGQGEPAVRVCHSL---VKHSQSRPSSRWROKIRTRTGGGPGS 98

RESULT 3
US-10-273-573-8476
; Sequence 8476, Application US/10273573
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: US/10/273,573
; CURRENT FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO: 8476
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(149)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-273-573-8476

Query Match
Best Local Similarity 38.9%; Score 56; DB 6; Length 149;
Matches 21; Conservative 4; Mismatches 15; Indels 14; Gaps 4;

Qy 4 GPGGDSPLQRPQHLMD--QGMRHSFSAPELLRQDKRPR---SGSTGS 49
Db 8 GNTGQRVPLPPHPPPIHVSRRKGLRHGF-----LKPMPERKLGESGKTGS 55

RESULT 4
US-10-273-573-9052
; Sequence 9052, Application US/10273573
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: US/10/273,573
; CURRENT FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO: 9052
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
```

```
; NAME/KEY: DOMAIN
; LOCATION: (176)..(202)
; OTHER INFORMATION: pCIC-type peptidyl-prolyl cis-trans isomerase proteins domain
; OTHER INFORMATION: identified by eMATRIX, accession number B01096C, p-value=6.063e-
US-10-273-573-9052

Query Match
Best Local Similarity 36.6%; Score 56; DB 6; Length 205;
Matches 15; Conservative 5; Mismatches 17; Indels 4; Gaps 1;

Qy 4 GPGGDSPLQRPQHLMDQGMRHSFSAPELLRQDKRPR 44
Db 45 GKGQGEPAVRVCHSL---VKHSQSRPSSRWROKIRTRT 81

RESULT 5
US-09-291-417D-13
; Sequence 13, Application US/09291417D
; GENERAL INFORMATION:
; APPLICANT: PLUMMAN, GREGORY
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHITE, DAVID
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
; FILE REFERENCE: 038602/0329
; CURRENT APPLICATION NUMBER: US/09/291,417D
; CURRENT FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 60/081,784
; PRIOR FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO: 13
; LENGTH: 1239
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-291-417D-13

Query Match
Best Local Similarity 40.0%; Score 52; DB 5; Length 1239;
Matches 20; Conservative 4; Mismatches 15; Indels 6; Gaps 3;

Qy 8 QDSPLQRPQHLMDQGMRHSFSAPELL--RODK---RPRSGST 47
Db 637 RDSPLQSGGQNSQAGQ--RNSTSIPEPLMERVELVPRPSGSS 680

RESULT 6
US-10-273-573-7947
; Sequence 7947, Application US/10273573
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: US/10/273,573
; CURRENT FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO: 7947
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(139)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-273-573-7947

Query Match
Best Local Similarity 37.5%; Score 51; DB 6; Length 139;
Matches 19; Conservative 5; Mismatches 17; Indels 4; Gaps 1;
```

Matches 21; Conservative 1; Mismatches 22; Indels 12; Gaps 4;
Qy 2 PLGPR---GDSPLQ-RPQHLMDQGMHRSFSGAG---PELLRQDRPRSGSTG 48
Db 78 PPGDVMGSGPPTPLQHRSPH---PGBERHGFSSKXPPPPGSPQKERRKQCG 130

RESULT 7
US-10-273-573-8693
; Sequence 8693, Application US/10273573
; GENERAL INFORMATION:
; APPLICANT: Hysq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: US/10/273,573
; CURRENT FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 8693
; LENGTH: 167
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(167)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-273-573-8693

Query Match 19.4%; Score 50.5; DB 6; Length 167;
Best Local Similarity 31.7%; Pred. No. 3.1;
Matches 13; Conservative 5; Mismatches 10; Indels 13; Gaps 1;
Qy 2 PLGPRGDSPLQRPQHLMDQGMHRSFSGAPPELLRQDRP 42
Db 113 PPGGGEAPPLR-----SSGPRAPLRPSRSP 140

RESULT 8
US-09-291-417D-89
; Sequence 89, Application US/09291417D
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHITE, DAVID
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
; FILE REFERENCE: 038602/0329
; CURRENT APPLICATION NUMBER: US/09/291,417D
; CURRENT FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 60/081,784
; PRIOR FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 89
; LENGTH: 1233
; TYPE: PRT
; ORGANISM: Murine sp.
US-09-291-417D-89

Query Match 19.4%; Score 50.5; DB 5; Length 1233;
Best Local Similarity 40.0%; Pred. No. 35;
Matches 18; Conservative 4; Mismatches 18; Indels 5; Gaps 2;

Qy 8 QDSPLQRPQHLMDQGMHRSFSGAPPELL--RQDK--RPRSGST 47
Db 627 RDSPLQGGQGNQAGQNRSTSIIPRLMERVELVPRPGSGSS 671

RESULT 9
US-10-273-573-7079

; Sequence 7079, Application US/10273573
; GENERAL INFORMATION:
; APPLICANT: Hysq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: US/10/273,573
; CURRENT FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 7079
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (463)..(486)
; OTHER INFORMATION: TRANSFORMING PROTEIN P21 RAS SIGNATURE domain identified by Pfam, accession name ras, E-
; OTHER INFORMATION: EMATRX, accession number PR00449E, p-value=8.714e-13, raw score
; OTHER INFORMATION: 13.50
; OTHER INFORMATION: value=0.0015, Pfam score of -96.4
US-10-273-573-7079

Query Match 19.2%; Score 50; DB 6; Length 501;
Best Local Similarity 26.8%; Pred. No. 14;
Matches 15; Conservative 10; Mismatches 13; Indels 18; Gaps 2;

Qy 3 LGPRGDSPLQRPQH-----LMDQGMHRSFSGAPPELLRQDK 40
Db 281 IGPRGKIGQKKNRQHGSKSVTVRLAVTKTRPVNIOIEFTTETLTGKDLRDK 336

RESULT 10
US-10-273-573-7748
; Sequence 7748, Application US/10273573
; GENERAL INFORMATION:
; APPLICANT: Hysq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: US/10/273,573
; CURRENT FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 7748
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (463)..(486)
; OTHER INFORMATION: TRANSFORMING PROTEIN P21 RAS SIGNATURE domain identified by Pfam, accession name ras, E-
; OTHER INFORMATION: EMATRX, accession number PR00449E, p-value=8.714e-13, raw score
; OTHER INFORMATION: 13.50
; OTHER INFORMATION: value=0.0015, Pfam score of -96.4
US-10-273-573-7748

Query Match 19.2%; Score 50; DB 6; Length 501;
Best Local Similarity 26.8%; Pred. No. 14;

Matches 15; Conservative 10; Mismatches 13; Indels 18; Gaps 2;

QY 3 LGPRGDSPLLRPOH-----LMDGGMRSFSAGPEL-LRQDK 40

DB 281 IGPKGKIGQKNRTOHGSKSVTRVLRAVTKTRPVNQIEFRFTTFLGTLDLRDCK 336

RESULT 11

US-10-273-573-8445

; Sequence 8445, Application US/10273573

; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc

; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES

; FILE REFERENCE: 21272-066

; CURRENT APPLICATION NUMBER: US/10/273,573

; PRIOR FILING DATE: 2002-10-18

; PRIOR APPLICATION NUMBER: 09/522,929

; PRIOR FILING DATE: 2000-04-18

; PRIOR APPLICATION NUMBER: 09/770,160

; NUMBER OF SEQ ID NOS: 10994

; SOFTWARE: Custom

; SEQ ID NO 8445

; LENGTH: 501

; TYPE: PRT

; ORGANISM: Homo sapiens

; NAME/KEY: DOMAIN

; LOCATION: (463)..(486)

; OTHER INFORMATION: TRANSFORMING PROTEIN P21 RAS SIGNATURE domain identified by

; OTHER INFORMATION: EMATRIX, accession number PR00449E, p-value=8.714e-13, raw score

; FEATURE: 13.50

; NAME/KEY: DOMAIN

; LOCATION: (287)..(501)

; OTHER INFORMATION: Ras family domain identified by Pfam, accession name ras, E-

; OTHER INFORMATION: value=0.0015, Pfam score of -96.4

US-10-273-573-8445

Query Match 19.2%; Score 50; DB 6; Length 501;

Best Local Similarity 26.8%; Pred. No. 14;

Matches 15; Conservative 10; Mismatches 13; Indels 18; Gaps 2;

QY 3 LGPRGDSPLLRPOH-----LMDGGMRSFSAGPEL-LRQDK 40

DB 281 IGPKGKIGQKNRTOHGSKSVTRVLRAVTKTRPVNQIEFRFTTFLGTLDLRDCK 336

RESULT 12

PCT-US02-18638A-186

; Sequence 186, Application PC/TUS0218638A

; GENERAL INFORMATION:

; APPLICANT: Millennium Pharmaceuticals, Inc. et al.

; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR

; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY

; FILE REFERENCE: MRI-035PC

; CURRENT APPLICATION NUMBER: PCT/US02/18638A

; PRIOR FILING DATE: 2002-06-12

; PRIOR APPLICATION NUMBER: US 60/298,159

; PRIOR FILING DATE: 2001-06-13

; PRIOR APPLICATION NUMBER: US 60/298,155

; PRIOR FILING DATE: 2001-06-13

; PRIOR APPLICATION NUMBER: US 60/335,936

; PRIOR FILING DATE: 2001-11-14

; NUMBER OF SEQ ID NOS: 238

; SOFTWARE: FaastSeq for Windows Version 4.0

; SEQ ID NO 186

; LENGTH: 603

; TYPE: PRT

; ORGANISM: Homo sapiens

PCT-US02-18638A-186

Query Match 19.2%; Score 50; DB 1; Length 603;

Best Local Similarity 27.3%; Pred. No. 17;

Matches 15; Conservative 9; Mismatches 11; Indels 20; Gaps 2;

QY 6 RGDSPLLRPO-----HMDGGMRSFSAGPE---LRQDK 40

DB 345 KGLNPLPERPREKEPVVRETGEVVDCHLSDMLQQLHSVNAKSEKGLVRQEE 399

RESULT 13

US-10-273-573-8772

; Sequence 8772, Application US/10273573

; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc

; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES

; FILE REFERENCE: 21272-066

; CURRENT APPLICATION NUMBER: US/10/273,573

; PRIOR FILING DATE: 2002-10-18

; PRIOR APPLICATION NUMBER: 09/522,929

; PRIOR FILING DATE: 2000-04-18

; PRIOR APPLICATION NUMBER: 09/770,160

; NUMBER OF SEQ ID NOS: 10994

; SOFTWARE: Custom

; SEQ ID NO 8772

; LENGTH: 87

; TYPE: PRT

; ORGANISM: Homo sapiens

; NAME/KEY: misc feature

; LOCATION: (1)..(87)

; OTHER INFORMATION: Xaa = X or * as defined in Table 2

US-10-273-573-8772

Query Match 18.8%; Score 49; DB 6; Length 87;

Best Local Similarity 27.1%; Pred. No. 23;

Matches 16; Conservative 5; Mismatches 26; Indels 12; Gaps 1;

QY 3 LGPRGDSPLLRPOHMDGGMRS-----FSAGPELRLQDKRRSGSTGS 49

DB 8 LGSLOPPPPVLRSSHLSPSSWEYSHTCNFCIFCRDGFVAVLPRLLGXSNRPASSXNT 66

RESULT 14

US-10-464-368-49

; Sequence 49, Application US/10464368

; GENERAL INFORMATION:

; APPLICANT: Krumlauf, Robb

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REGULATING BONE DEPOSITION

; FILE REFERENCE: 40716-IP-017

; CURRENT APPLICATION NUMBER: US/10/464,368

; PRIOR FILING DATE: 2003-06-16

; PRIOR APPLICATION NUMBER: 60/388,970

; PRIOR FILING DATE: 2002-06-14

; NUMBER OF SEQ ID NOS: 140

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 49

; LENGTH: 205

; TYPE: PRT

; ORGANISM: MOUSE

US-10-464-368-49

Query Match 18.8%; Score 49; DB 6; Length 205;

Best Local Similarity 28.8%; Pred. No. 6.4;

Matches 15; Conservative 10; Mismatches 17; Indels 10; Gaps 3;

QY 5 PRGDSPL-----ORPGLMDGGMRSFSAGPELRLQDK--RPRSGSTGS 49

DB 152 PSRRVRLVASCCKRPTRFNQSELK--DFGPETARPOXGRKRPGARQA 200

RESULT 15

```
US-10-273-573-7365
; Sequence 7365, Application US/10273573
; GENERAL INFORMATION:
; APPLICANT: Hyseg, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: US/10/273,573
; CURRENT FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 7365
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-273-573-7365

Query Match      18.8% Score 49; DB 6; Length 364;
Best Local Similarity 25.9% Pred. No. 13;
Matches 15; Conservative 9; Mismatches 14; Indels 20; Gaps 2;

Oy      10 SPLLRPQHLMDDCGMRHSF---SAGPELRLRQDKR-----PRSGST 47
      ||| | | | | | | | | | | | | | | | | | | | | | | | | |
Db      43 SPLLGPRPILRQGTILETLCLPTSAPELHRSRRSDAKCSSAAVVRGSALEPRPGT 100
```

Search completed: July 24, 2003, 12:16:54
Job time : 1.63485 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 11:48:38 ; Search time 4.01757 Seconds
(without alignments)
1172.914 Million cell updates/sec

Title: US-09-991-681-31

Perfect score: 260
Sequence: 1 EPLGPRGQDSPILQRPQHLM.....SAGPELLRQDKRPRSGTGS 49

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	63	24.2	462	2	G83000
2	62.5	24.0	1655	2	T13998
3	60	23.1	333	2	E75332
4	59.5	22.9	345	2	D72536
5	59	22.7	271	2	S30230
6	58.5	22.5	276	2	S54830
7	58.5	22.5	3429	2	T13853
8	58	22.3	270	2	JN0767
9	58	22.3	357	2	H87433
10	57.5	22.1	405	2	A02784
11	57.5	22.1	405	2	D97563
12	57	21.9	533	2	S69336
13	57	21.9	536	2	D83622
14	56.5	21.7	313	1	YXHT
15	56.5	21.7	807	2	D86350
16	56	21.5	163	2	S68520
17	56	21.5	277	2	S78063
18	56	21.3	1329	2	T29074
19	55.5	21.3	385	2	G87519
20	55.5	21.3	455	2	AH0794
21	55.5	21.3	1615	2	U65510
22	55	21.2	93	2	A06740
23	55	21.2	186	1	RSBY8E
24	55	21.2	280	2	T03543
25	55	21.2	416	2	T34279
26	54.5	21.0	103	2	A72713
27	54.5	21.0	170	1	NMMU2
28	54.5	21.0	229	2	B72700
29	54.5	21.0	848	2	S02262

30	54.5	21.0	881	2	G96574	hypothetical prote
31	54.5	21.0	1433	2	G01946	nitric-oxide synth
32	54	20.8	165	2	JC7136	peptidylprolyl iso
33	54	20.8	284	2	T35723	cobalt transport s
34	54	20.8	473	2	S09775	hypothetical prote
35	54	20.8	623	1	S3167	gene pointed prote
36	53.5	20.6	2212	2	A41098	calcium channel pr
37	53.5	20.6	463	2	AB3521	replication protei
38	53.5	20.6	1675	1	LRTH	probable histone a
39	53.5	20.6	3063	2	J50156	clathrin heavy cha
40	53	20.4	561	2	D96674	genome polyprotein
41	53	20.4	938	2	T34105	hypothetical prote
42	53	20.4	1527	2	H70655	hypothetical prote
43	53	20.4	2715	2	T13049	glutamate synthase
44	53	20.4	184	2	F75481	eyelid - fruit fly
45	52.5	20.2				hypothetical prote

ALIGNMENTS

RESULT 1
G83000
Probable two-component response regulator PA5166 [imported] - Pseudomonas aeruginosa (str
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 17-Mar-2003
C:Accession: G83000
R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.V.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapid, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: AB2950; MUID:20437337; PMID:10984043
A:Accession: G83000
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-462 <STO>
A:Cross-references: GB:AE004929; GB:AE004091; NID:9951462; PIDN:AG08551.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA5166
C:Superfamily: response regulator of the NtrC type; response regulator homology; RNA pol

Query Match 24.2%; Score 63; DB 2; Length 462;
Best Local Similarity 48.3%; Pred. No. 3.6;
Matches 14; Conservative 3; Mismatches 10; Indels 2; Gaps 1;

OY 2 PLAGRQDSPILQRPQHLMDGQNRHSFS 30
DB 322 PLRREGDIPILFR--HFAEGAMRHGILT 348

RESULT 2
T13998
gene mastermind protein - fruit fly (Drosophila virilis)
C:Species: Drosophila virilis
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Nov-2000
C:Accession: T13998
R:Newfeld, S.V.; Tachida, H.; Yedvobnick, B.
J. Mol. Evol. 38, 637-641, 1994
A:Title: Drive-selection equilibrium: homopolymer evolution in the Drosophila gene maste
A:Reference number: Z17850; MUID:94365848; PMID:8083889
A:Accession: T13998
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-1655 <NEW>
A:Cross-references: EMBL:M92914; NID:9157833; PID:9157834; PIDN:AAC37201.1
C:Genetics:
A:Cross-references: FlyBase:FBgn003119

Query Match 24.0%; Score 62.5; DB 2; Length 1655;
Best Local Similarity 27.6%; Pred. No. 17;
Matches 16; Conservative 9; Mismatches 22; Indels 11; Gaps 1;

QY 2 PLGPRGDSPLQRPQHLMDOGMKHSFSAG-----PELLRDKRRRSGSTG 48
 DB 1451 PNGPQMQLTPAQMQQQHMKQDQDQHMGPBGGGGGGGMQMLLQDQDQNAAGGGG 1508

RESULT 3

E75332
 Probable ADP-ribosylglycohydrolase - Deinetococcus radiodurans (strain R1)
 C:Species: Deinetococcus radiodurans
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
 C:Accession: E75332
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium Deinetococcus radiodurans R1.
 A:Reference number: A75250; MUID:20036896; PMID:10567266
 A:Accession: E75332
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-333 <WHI>
 A:Cross-references: GB:AE002034; GB:AE000513; NID:g6459742; PIDN:AAF11509.1; PID:g645974
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DR1958
 A:Map position: 1

Query Match 23.1%; Score 60; DB 2; Length 333;
 Best Local Similarity 34.9%; Pred. No. 5.9;
 Matches 15; Conservative 7; Mismatches 11; Indels 10; Gaps 1;

QY 4 GRRGDSPLQRPQHLMDOGMKHSFSAG-----PELL 36
 DB 230 GARGTSPLEVEPEHLPELRLGHGWPQWRDWAETHPPELL 332

RESULT 4

D72536
 Hypothetical protein APE1577 - Aeropyrum pernix (strain K1)
 C:Species: Aeropyrum pernix
 C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
 C:Accession: D72536
 R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahara, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, T.; KDNA Res. 6, 83-101, 1999
 A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix strain K1.
 A:Reference number: A72450; MUID:99310339; PMID:10382966
 A:Accession: D72536
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-345 <KAW>
 A:Cross-references: DDBJ:AP000062; NID:g5105244; PIDN:BAA80577.1; PID:dl044363; PID:g510
 A:Experimental source: strain K1
 C:Genetics:
 A:Gene: APE1577

Query Match 22.9%; Score 59.5; DB 2; Length 345;
 Best Local Similarity 31.9%; Pred. No. 7.1;
 Matches 23; Conservative 8; Mismatches 16; Indels 25; Gaps 5;

QY 1 EPLG-----PRGDSPLQRPQHLMDO-----GGM-----RHSFSAGPEL-----LR 37
 DB 244 EPLGHHRYELGAPPGGEPHPEVLNFDLATSPGALGLCLHPHAAVSPPSLSQHAHR 303

QY 38 --QDKRPRSGST 47
 DB 304 WSPTRKRGSS 315

RESULT 5

S30230
 homeotic protein Hex - mouse

C:Species: Mus musculus (house mouse)
 C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 24-Sep-1999
 C:Accession: S30230
 R:Bedford, F.K.; Ashworth, A.; Enver, T.; Wiedemann, L.M.
 Nucleic Acids Res. 21, 1245-1249, 1993
 A:Title: HEX: a novel homeobox gene expressed during haematopoiesis and conserved between
 A:Reference number: S30230; MUID:99219088; PMID:8096636
 A:Accession: S30230
 A:Molecule type: DNA
 A:Residues: 1-271 <BED>
 A:Cross-references: EMBL:Z21524; NID:g288500; PIDN:CAA79729.1; PID:g288501
 C:Superfamily: unassigned homeobox proteins; homeobox homology
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F:139-195/Domain: homeobox homology <HOX>

Query Match 22.7%; Score 59; DB 2; Length 271;
 Best Local Similarity 45.7%; Pred. No. 6.2;
 Matches 16; Conservative 2; Mismatches 17; Indels 0; Gaps 0;

QY 1 EPLGPRGDSPLQRPQHLMDOGMKHSFSAGPEL 35
 DB 119 DPLGRLMSPFLQRPKHKGQVRFSDQTVEL 153

RESULT 6

S54830
 P1 protein - potato virus Y (fragment)
 C:Species: potato virus Y, PVY
 C:Date: 08-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 17-Nov-2000
 C:Accession: S54830
 R:Pehu, T.; Maki-Valkama, T.; Valkonen, J.P.T.; Koivu, K.; Lehto, K.; Pehu, E.
 submitted to the EMBL Data Library, November 1994
 A:Description: Potato plants transformed with a potato virus Y P1 gene sequence are resis
 A:Reference number: S54830
 A:Accession: S54830
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-276 <PEH>
 A:Cross-references: EMBL:X82848; NID:g817817; PIDN:CAA58051.1; PID:g817818
 C:Superfamily: tobacco etch virus genome polyprotein

Query Match 22.5%; Score 58.5; DB 2; Length 276;
 Best Local Similarity 34.5%; Pred. No. 7.4;
 Matches 19; Conservative 8; Mismatches 15; Indels 13; Gaps 3;

QY 1 EPLGPRGDSPLQ-----RPQHLMDOGMKHSFSAGPELDRKRRSGSTG 49
 DB 117 EPQAPRGIIHTTPRVKVKTRPIIKLTGQMDH-----LIKQVKQIMSGKRG 164

RESULT 7

T13853
 Hypothetical protein X - fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Nov-2000
 C:Accession: T13853
 R:Kraemer, C.; Schmidt, E.R.
 submitted to the EMBL Data Library, April 1996
 A:Description: A novel gene, DmX, on the Drosophila X-chromosome; a homologue to Cpy fro
 A:Reference number: Z17797
 A:Accession: T13853
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-3429 <KRA>
 A:Cross-references: EMBL:X97196; NID:g1279383; PID:e235567; PIDN:CAA65830.1
 C:Genetics:
 A:Gene: DMX
 A:Cross-references: FlyBase:FBgn0023458
 A:Map position: X
 A:Introns: 29/3; 102/2; 152/1; 2478/1; 2838/1; 2913/1; 3032/1; 3120/2; 3149/3; 3239/3; 3

Query Match 22.5%; Score 58.5; DB 2; Length 3429;
 Best Local Similarity 41.0%; Pred. No. 1.3e+02;

Matches 16; Conservative 5; Mismatches 11; Indels 7; Gaps 2;

Ox 9 DSPILQRPQHLMDOGQMRHSFSACPELLLRQDKRRPSGST 47

Db 1334 DSRLAKKQRL---NQVLHTEPAHDPE---SRSPSGST 1365

RESULT 8

UN0767
homeobox protein HEX - human
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 20-Aug-1994 #text_change 24-Sep-1999
C:Accession: JN0767; S26799; S78048; S30231
R:Herman, R.; Radich, J.; Collins, S.
Biochem. Biophys. Res. Commun. 195, 976-983, 1993
A:Title: PCR cloning of an orphan homeobox gene (PRH) preferentially expressed in myeloid
A:Reference number: JN0767; MUID:93384629; PMID:8103988
A:Accession: JN0767
A:Molecule type: DNA
A:Residues: 1-270 <ACC>
A:Cross-references: GB:LI6499; NID:g3292404; PIDN:AA02988.1; PID:g292405
R:Crompton, M.R.; Bartlett, T.J.; MacGregor, A.D.; Manfioletti, G.; Baurati, E.; Giancotti
Nucleic Acids Res. 20, 5661-5667, 1992
A:Title: Identification of a novel vertebrate homeobox gene expressed in haematopoietic
A:Reference number: S26799; MUID:93087175; PMID:1360645
A:Accession: S26799
A:Molecule type: mRNA
A:Residues: 1-114, 'L', 116-270 <CRO>
A:Cross-references: EMBL:X67235; NID:g32547; PIDN:CAA47661.1; PID:g32548
R:Bedford, F.K.
Submitted to the EMBL Data Library, February 1993
A:Reference number: S78048
A:Accession: S78048
A:Molecule type: DNA
A:Residues: 122-270 <BED>
A:Cross-references: EMBL:Z21533; NID:g32068; PIDN:CAA79730.1; PID:g32069
R:Bedford, F.K.; Ashworth, A.; Enver, T.; Wiedemann, L.M.
Nucleic Acids Res. 21, 1245-1249, 1993
A:Title: HEX: a novel homeobox gene expressed during haematopoiesis and conserved between
A:Reference number: S30230; MUID:93219088; PMID:8096636
A:Accession: S30231
A:Molecule type: DNA
A:Residues: 136-195 <BEW>
A:Cross-references: EMBL:Z21533
C:Comment: This protein is an important regulator of normal hematopoiesis.
C:Genetics:
A:Gene: HEX
C:Superfamily: unassigned homeobox proteins; homeobox homology
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
F:138-194/Domain: homeobox homology <HOX>

Query Match 22.3%; Score 58; DB 2; Length 270;
Best Local Similarity 50.0%; Pred. No. 8.3; Mismatches 12; Indels 0; Gaps 0;
Matches 14; Conservative 2; Mismatches 12; Indels 0; Gaps 0;

Ox 1 EPLGPRGQDSPILQRPQHLMDOGQMRHS 28

Db 118 DPLGKPLMSPLQRPPLKRRGGQVRFS 145

RESULT 9

H87433
transcription regulator, lacI family [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
C:Accession: H87433
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Bisen, J.; Heidelberg, J.
B.; Jamb, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolor
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001.
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: H87433

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-357 <STO>
A:Cross-references: GB:AE005673; NID:g13422862; PIDN:AAK23468.1; GSPDB:GN00148
C:Genetics:
A:Gene: CCI489
C:Superfamily: lac repressor

Query Match 22.3%; Score 58; DB 2; Length 357;
Best Local Similarity 34.4%; Pred. No. 11;
Matches 11; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

Ox 14 QPQHLMDOGMRHSFSAPELLRODKPRSG 45
Db 191 QAARHVELGHKRIAFISGPETFRSSHERRG 222

RESULT 10

AD2784
molybdopterin biosynthesis protein [imported] - Agrobacterium tumefaciens (strain C58, D
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C:Accession: AD2784
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavina, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AD2784
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-405 <KUB>
A:Cross-references: GB:AE008688; PIDN:AAI42690.1; PID:g17740125; GSPDB:GN00186
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: moea
A:Map position: circular chromosome
C:Superfamily: molybdenum cofactor biosynthesis protein moea-2

Query Match 22.1%; Score 57.5; DB 2; Length 405;
Best Local Similarity 36.6%; Pred. No. 15;
Matches 15; Conservative 8; Mismatches 11; Indels 7; Gaps 2;

Ox 5 PRGDSPLQRPQHLMDOGMRHSF--SAGPELLRODKPR 43
Db 103 PDGADTILIQEDAEILDGDKIRTFPAVTAGRAI-----RPR 138

RESULT 11

D97563
hypothetical protein AGR_C3105 [imported] - Agrobacterium tumefaciens (strain C58, Cere
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
C:Accession: D97563
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Mollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: D97563
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-405 <KUB>
A:Cross-references: GB:AE007869; PIDN:AAK87461.1; PID:g15156781; GSPDB:GN00169
C:Genetics:
A:Gene: AGR_C3105
A:Map position: circular chromosome
C:Superfamily: molybdenum cofactor biosynthesis protein moea-2

Query Match 22.1%; Score 57.5; DB 2; Length 405;

Best Local Similarity 36.6%; Pred. No. 15;
Matches 15; Conservative 8; Mismatches 11; Indels 7; Gaps 2;

Qy 5 PRGQSPPLQRPQHLMDQGMRSF--SAGPELLRQDKRPR 43
Db 103 PGADRTILIQEDAEILDDGKIRTFVAVTAGRH-----RPR 138

RESULT 12

S69336
Arylsulfatase (EC 3.1.6.1) - *Pseudomonas aeruginosa*

C:Species: *Pseudomonas aeruginosa*

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 17-Mar-2000

C:Accession: S69336; S69359

R:Belk, S.; Kehrl, H.; James, P.; Staudemann, W.; Cook, A.M.; Leisinger, T.; Kertesz, E.; J. Biochem. 229, 385-394, 1995

A:Title: Purification and characterization of the arylsulfatase synthesized by *Pseudomonas*

A:Reference number: S69336; PMID:95262702; PMID:7744061

A:Accession: S69336

A:Molecule type: DNA

A:Residues: 1-533 <BE>

A:Cross-references: EMBL:248540; NID:g2440146; PIDN:CAA68421.1; PID:S695684

A:Experimental source: strain PAO1

A:Accession: S69359

A:Molecule type: protein

A:Residues: 2-27;95-101;131-155;159-169;188-195;245-263;297-302;340-350;383-388;407-419;

R:Dieks, T.; Meich, C.; Hummerjahn, J.; Schmidt, B.; Kertesz, M.A.; von Figura, K.

J. Biol. Chem. 273, 25560-25564, 1998

A:Title: Posttranslational formation of formylglycine in prokaryotic sulfatases by modification

A:Reference number: A59074; PMID:98421466; PMID:9748219

A:Contents: annotation: post-translational modification

A>Note: confirmation of 3-oxoalanine active site, referred to as formylglycine

C:Genetics:

A:Gene: atsA

C:Superfamily: animal sulfatase

C:Keywords: sulfuric ester hydrolase

F:51/Modified site: 3-oxoalanine (Cys) #status experimental

Query Match 21.9%; Score 57; DB 2; Length 533;

Best Local Similarity 36.1%; Pred. No. 24;

Matches 13; Conservative 8; Mismatches 13; Indels 2; Gaps 1;

Qy 13 LORPQHLMDQGMRSFSGPELLRQDKRPRSGSTG 48
Db 214 LQAPREIVE--KYRGKYDAGPEALRQERLARKELG 247

RESULT 13

D83622
Arylsulfatase PA0183 [imported] - *Pseudomonas aeruginosa* (strain PAO1)

C:Species: *Pseudomonas aeruginosa*

C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C:Accession: D83622

R:Stover, C.K.; Pham, X.Q.; Ewlin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kass, A.; Laidig, K.; Lam,

.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic patho

A:Reference number: A82950; PMID:20437337; PMID:10984043

A:Accession: D83622

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-536 <STO>

A:Cross-references: GB:AE004456; GB:AE004091; NID:g9946013; PIDN:AAG03573.1; GSPDB:GN001

A:Experimental source: strain PAO1

C:Genetics:

A:Gene: atsA; PA0183

C:Superfamily: animal sulfatase

Query Match 21.9%; Score 57; DB 2; Length 536;

Best Local Similarity 36.1%; Pred. No. 24;

Matches 13; Conservative 8; Mismatches 13; Indels 2; Gaps 1;

Qy 13 LORPQHLMDQGMRSFSGPELLRQDKRPRSGSTG 48
Db 214 LQAPREIVE--KYRGKYDAGPEALRQERLARKELG 247

RESULT 14

YXHT

thymidylate synthase (EC 2.1.1.45) - human

C:Species: *Homo sapiens* (man)

C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 16-Jun-2000

C:Accession: A23047; I55318; J00120; A22393; A33842

R:Takeishi, K.; Kaneda, S.; Ayusawa, D.; Shimizu, K.; Gotch, O.; Sano, T.

Nucleic Acids Res. 13, 2035-2043, 1985

A:Title: Nucleotide sequence of a functional cDNA for human thymidylate synthase.

A:Reference number: A23047; PMID:85215597; PMID:2987839

A:Accession: A23047

A:Molecule type: mRNA

A:Residues: 1-313 <TA>

A:Cross-references: EMBL:X02308; NID:g37478; PIDN:CAA26178.1; PID:g37479

U. Kaneda, S.; Nalbantoglu, U.; Takeishi, K.; Shimizu, K.; Gotch, O.; Sano, T.; Ayusawa,

J. Biol. Chem. 265, 20277-20284, 1990

A:Title: Structural and Functional Analysis of the Human Thymidylate Synthase Gene.

A:Reference number: I55318; PMID:91056070; PMID:2243092

A:Accession: I55318

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-313 <RES>

A:Cross-references: GB:D00596; NID:g220135; PIDN:BA00472.1; PID:g220136

R:Takeishi, K.; Kaneda, S.; Ayusawa, D.; Shimizu, K.; Gotch, O.; Sano, T.

J. Biochem. 106, 575-583, 1989

A:Title: Human thymidylate synthase gene: isolation of phage clones which cover a functi

A:Reference number: J00120; PMID:9010051; PMID:2532645

A:Accession: J00120

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-68 <TA>

A:Cross-references: GB:D00517; NID:g220133; PIDN:BA00404.1; PID:g2160415

R:Shimizu, K.; Ayusawa, D.; Takeishi, K.; Sano, T.

J. Biochem. 97, 845-850, 1985

A:Title: Purification and NH₂-terminal amino acid sequence of human thymidylate synthase

A:Reference number: A22393; PMID:85261174; PMID:3839505

A:Accession: A22393

A:Molecule type: protein

A:Residues: 2-25 <SH>

R:Davidson, V.O.; Sitawaraporn, W.; Santi, D.V.

J. Biol. Chem. 264, 9145-9148, 1989

A:Title: Expression of human thymidylate synthase in *Escherichia coli*.

A:Reference number: A33842; PMID:89255401; PMID:2656695

A:Accession: A33842

A:Molecule type: protein

A:Residues: 2-10 <DAV>

C:Genetics:

A:Gene: GDB:TYMS

A:Cross-references: GDB:120465; OMIM:188350

A:Map position: 18p11.32-18p11.32

A:Introns: 69/1; 93/3; 152/1; 186/1; 244/3; 268/3

C:Superfamily: thymidylate synthase; thymidylate synthase homology

C:Keywords: deoxyribonucleotide biosynthesis; methyltransferase

F:30-313/Domains: thymidylate synthase homology <TDS>

F:199/Active site: Cys #status predicted

Query Match 21.7%; Score 56.5; DB 1; Length 313;

Best Local Similarity 34.0%; Pred. No. 15;

Matches 16; Conservative 7; Mismatches 19; Indels 5; Gaps 2;

Qy 2 PLGPRGDSPLQRPQHLMDQGMRSFSGPELLRQDKRPRSGS 46
Db 12 PLGPRGDSPLQRPQHLMDQGMRSFSGPELLRQDKRPRSGS 46

RESULT 15
D86350
F6K7.13 protein - *Arabidopsis thaliana*

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 11:46:57 ; Search time 2.22401 Seconds
(without alignments)
1036.105 Million cell updates/sec

Title: US-09-991-681-31

Perfect score: 260
Sequence: 1 EPLGPRGQDSPLQRPHLM.....SAGPELLRQDKRPRSGSTGS 49

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	61	23.5	250	1 HXB9_HUMAN	P17482 homo sapien
2	59.5	22.9	2004	1 MOZ_HUMAN	Q97794 homo sapien
3	58	22.7	271	1 HMPH_MOUSE	P43120 mus musculu
4	58	22.3	270	1 HMPH_HUMAN	Q03014 homo sapien
5	58	22.3	943	1 NFEC1_HUMAN	O95644 homo sapien
6	57	21.9	485	1 ONC2_HUMAN	O95948 homo sapien
7	57	21.9	535	1 ARS_PSEAE	P51691 pseudomonas
8	56.5	21.7	312	1 TTSY_HUMAN	P04818 homo sapien
9	56.5	21.7	431	1 SP7_HUMAN	Q13526 homo sapien
10	56	21.5	163	1 PIN1_HUMAN	Q15502 gallus gall
11	56	21.5	277	1 HMPH_CHICK	P37418 salmonella
12	55.5	21.3	455	1 MENE_SALTY	Q9UJ99 homo sapien
13	55.5	21.3	622	1 ARS_HUMAN	P07279 saccharomyc
14	55	21.2	186	1 RL18_YEAST	Q95244 homo sapien
15	55	21.2	1629	1 ARS9_HUMAN	P15458 arabidopsis
16	54.5	21.0	170	1 2SS2_ARATH	Q96477 mus musculu
17	54	20.8	165	1 PIN1_MOUSE	P16755 human cytom
18	54	20.8	473	1 UH13_HCMVA	P51022 drosophila
19	54	20.8	623	1 PNT1_DROME	P54282 rattus norv
20	54	20.8	2212	1 CCA1_RAT	Q8V167 mus musculu
21	53.5	20.6	428	1 SP7_MOUSE	Q00610 homo sapien
22	53.5	20.6	1675	1 CLH1_HUMAN	P49951 bos taurus
23	53.5	20.6	1675	1 CLH1_BOVIN	P11442 rattus norv
24	53.5	20.6	1675	1 CLH1_RAT	P12422 p genome po
25	53.5	20.6	3063	1 POLG_PVN	Q96777 arabidopsis
26	53	20.4	561	1 4C13_ARATH	P48018 rattus norv
27	53	20.4	677	1 RGS9_RAT	Q96018 homo sapien
28	52.5	20.2	1336	1 W146_HUMAN	O95208 ovine aries
29	52.5	20.2	723	1 ICAL_SHEEP	P22602 porcato viru
30	52.5	20.2	856	1 POLG_PVVO	O91210 mus musculu
31	52.5	20.2	3301	1 CLAR3_MOUSE	O35885 mus musculu
32	52	20.0	263	1 ASH2_MOUSE	Q59653 pseudomonas
33	52	20.0	334	1 PYRB_PSEAE	

34	52	20.0	334	1 PYRB_PSEPU	Q59711 pseudomonas
35	52	20.0	675	1 RGS9_MOUSE	O54828 mus musculu
36	52	20.0	704	1 SYN1_RAT	P09951 rattus norv
37	52	20.0	706	1 SYN1_BOVIN	P17599 bos taurus
38	52	20.0	809	1 PBPA_XYLEA	O9P644 xylella fas
39	52	20.0	1239	1 M4K4_HUMAN	O55819 homo sapien
40	52	20.0	1640	1 CLH2_HUMAN	P53675 homo sapien
41	51.5	19.8	164	1 2SS3_ARATH	P15459 arabidopsis
42	51.5	19.8	275	1 VINT_FRG1V	P29164 frog virus
43	51.5	19.8	759	1 MAO2_ECOLI	P76558 escherichia
44	51.5	19.8	2832	1 NDVB_RHIME	P20471 rhizobium m
45	51	19.6	346	1 MGB4_HUMAN	O15461 homo sapien

ALIGNMENTS

RESULT 1
HXB9_HUMAN STANDARD; PRT; 250 AA.
AC P17482; O9H11;
DT 01-AUG-1990 (Rel. 15, Created)
DT 26-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Homeobox protein Hox-B9 (Hox-2E) (Hox-2.5).
GN HOXB9 OR HOX2E.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kosaki K., Kosaki R., Suzuki T., Yoshinashi H., Sasaki K., Matsuo N.;
RT "A complete mutation analysis panel of human HOX genes";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RX MEDLINE=22388257; PubMed=12477932;
RA Strauberg R.L., Pelngold E.A., Grouse L.H., Derge J.G.,
RA Klusner R.D., Collins F.S., Wagner L., Shermen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heich F.,
RA Ditchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedini T.B., Toehiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loguigliano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosnak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krzywicki M.I., Skalek U., Skallue D.E.,
RA Scherch A., Schein J.B., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RP SEQUENCE OF 173-250 FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=89378558; PubMed=2570724;
RA Giampolo A., Acampora D., Zappavigna V., Pannese M.,
RA D'Esposito M., Care A., Faiella A., Stornaiuolo A., Russo G.,
RA Simeone A., Boncinelli E., Peschle C.;
RT "Differential expression of human HOX-2 genes along the anterior-
RL posterior axis in embryonic central nervous system";
RN [4]
RP SEQUENCE OF 185-250 FROM N.A.
RX MEDLINE=9021525; PubMed=2576652;
RA Boncinelli E., Acampora D., Pannese M., D'Esposito M., Somma R.,

```

RA Gaudio G.; Stornaiuolo A., Caffiero M., Faiella A., Simeone A.;
RT "Organization of human class I homeobox genes.";
RL Genome 31:745-756(1989).
CC -1- FUNCTION: SEQUENCE-SPECIFIC TRANSCRIPTION FACTOR WHICH IS PART OF
CC A DEVELOPMENTAL REGULATORY SYSTEM THAT PROVIDES CELLS WITH
CC SPECIFIC POSITONAL IDENTITIES ON THE ANTERIOR-POSTERIOR AXIS.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN WHOLE EMBRYOS AND FETUSES AT
CC 5-9 WEEKS FROM CONCEPTION.
CC -1- SIMILARITY: BELONGS TO THE ABD-B HOMEBOX FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AY014296; AAG42144.1; -.
DR EMBL: AY014295; AAG42144.1; JOINED.
DR EMBL: BC015565; AAH15565.1; -.
DR EMBL: X16172; CAA34294.1; -.
DR PIR: A37042; A37042.
DR HSSP: P02833; 9ANT.
DR TRANSFAC: T01738; -.
DR Genew: HGNC:5120; HOXB9.
DR MIM: 142964; -.
DR GO: GO:0005634; C:nucleus; NAS.
DR GO: GO:0003700; P:transcription factor activity; NAS.
DR GO: GO:0007275; P:development; NAS.
DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
DR InterPro: IPR001356; Homeobox.
DR Pfam: PF00046; homeobox; 1.
DR Pfam: PF04617; Hox9 act; 1.
DR PRINTS: PR00024; HOMEBOX.
DR PRINTS: PR00031; HTHREPRESSR.
DR PRODOM: PD00010; Homeobox; 1.
DR SMART: SM00389; HOX; 1.
DR PROSITE: PS00027; HOMEBOX 1; 1.
DR PROSITE: PS00071; HOMEBOX 2; 1.
DR Homeobox; DNA-binding: Developmental protein; Nuclear protein;
KW Transcription regulation.
FT DNA_BIND 185 244 HOMEBOX.
FT CONFLICT 173 173 T -> A (IN REF. 3).
FT SEQUENCE 250 AA; 28058 MW; F06ECBC08FBED2C CRC64;
SQ
Query Match 23.5%; Score 61; DB 1; Length 250;
Best Local Similarity 34.5%; Pred. No. 2.4;
Matches 20; Conservative 7; Mismatches 17; Indels 14; Gaps 3;
QY 1 BELGRGQDS-----PLLRPHLMDQGMHSF--SAGEELLRQDKRRSG 45
DB 100 EF-ARGEAARPGQAAYVAEPLLAGRGLLKQGRPEYSLERSAGREAVLNSGRGCG 156

```

```

RX MEDLINE=96376968; PubMed=8782817;
RA Borrow J., Stanton V.P. Jr., Andresen J.M., Becher R., Behn F.G.,
RA Chaganti R.S.K., Clavin C.I., Distèche C., Dube I., Fritchaut A.M.,
RA Horzani D., Mittleman F., Volinia S., Watmore A.E., Housman D.E.;
RT "The translocation t(8;16)(p11;p13) of acute myeloid leukaemia fuses
RT a putative acetyltransferase to the CREB-binding protein.";
RL Nat. Genet. 14:33-41(1996).
CC -1- FUNCTION: MAY REPRESENT A CHROMATIN-ASSOCIATED ACETYLTRANSFERASE.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC DISEASE: PARTICIPATES IN A T(8;16)(p11;p13) CHROMOSOMAL
CC TRANSLOCATION THAT PRODUCES A MOZ-CBP CHIMERA OBSERVED IN THE
CC M4/M5 SUBTYPE OF ACUTE MYELOID LEUKEMIA.
CC -1- SIMILARITY: BELONGS TO THE MYST (SAS/MOZ) FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U47742; AAC50662.1; -.
DR Genew: HGNC:13013; RUNXBP2.
DR MIM: 601408; -.
DR GO: GO:0006323; P:DNA packaging; TAS.
DR InterPro: IPR005818; Histone H1/H5.
DR InterPro: IPR002717; MOZ SAS.
DR InterPro: IPR001965; Znf_PHD.
DR Pfam: PF01853; MOZ_SAS; 1.
DR Pfam: PF00628; PHD; 2.
DR SMART: SM00526; H15; 1.
DR SMART: SM00249; PHD; 2.
DR PROSITE: PS01359; ZF_PHD 1; 1.
DR PROSITE: PS00016; ZF_PHD 2; 2.
DR Proto-oncogene; Chromosomal translocation; zinc-finger; Repeat;
KW Nuclear protein.
FT ZN_FING 206 265 PHD-TYPE 1.
FT ZN_FING 259 313 PHD-TYPE 2.
FT DOMAIN 371 379 POLY-SER.
FT ZN_FING 538 560 C2HC-TYPE.
FT DOMAIN 788 801 POLY-GLU.
FT DOMAIN 989 995 POLY-GLU.
FT DOMAIN 1019 1026 POLY-ARG.
FT DOMAIN 1069 1078 POLY-GLU.
FT DOMAIN 1147 1150 POLY-LYS.
FT DOMAIN 1221 1242 GLU-RICH.
FT DOMAIN 1267 1302 GLU-RICH.
FT DOMAIN 1411 1414 POLY-GLU.
FT DOMAIN 1593 1597 POLY-SER.
FT DOMAIN 1643 1704 GLN/PRO-RICH.
FT DOMAIN 1897 1977 MET-RICH.
FT SITE 1546 1547 BREAKPOINT FOR TRANSLOCATION TO FORM
FT MOZ-CBP.
SQ SEQUENCE 2004 AA; 225054 MW; 9FFBAC3792854BA CRC64;
Query Match 22.9%; Score 59.5; DB 1; Length 2004;
Best Local Similarity 32.1%; Pred. No. 41;
Matches 17; Conservative 9; Mismatches 16; Indels 11; Gaps 2;
QY 5 PRGQDSPILQRP-----QHLMDQGMHSFSAAGEELLRQDKRRSGSGTGS 49
DB 1679 PQQQPQPQPQPQPQPQPQPQPQPQPQPQPQPQPQPQPQPQPQPQPQPQPQP 1728

```

```

RESULT 2
MOZ_HUMAN
ID MOZ_HUMAN STANDARD; PRT; 2004 AA.
AC 092794;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Runct-related transcription factor binding protein 2 (Monocytic
DE leukemia zinc finger protein) (zinc finger protein 220).
GN RUNXBP2 OR ZNF220 OR MOZ.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCB1_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RESULT 3
HMPH_MOUSE
ID HMPH_MOUSE STANDARD; PRT; 271 AA.
AC P43120;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)

```

DE Homeobox protein PRH (hematopoietically expressed homeobox) (Homeobox
DE protein HHEX).
GN HHEX OR PRHX.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; Tissue=Embryo;
RX MEDLINE=93219088; PubMed=8096636;
RA Bedford F.K., Ashworth A., Enver T., Wiedemann L.M.;
"HEX: a novel homeobox gene expressed during haematopoiesis and
RT conserved between mouse and human."
RL Nucleic Acids Res. 21:1245-1249(1993).
CC -1- FUNCTION: RECOGNIZES THE DNA SEQUENCE 5'-ATTAA-3'. MAY PLAY A ROLE
CC IN HEMATOPOIETIC DIFFERENTIATION.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- DEVELOPMENTAL STAGE: EXPRESSED DURING HEMATOPOIESIS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; Z21524; CA79729.1; -.
DR PIR; S30230; S30230.
DR HSSP; P22808; INK3.
DR TRANSFAC; T03417; -.
DR MGD; MGI:96086; Hhex.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0007420; P:brain development; IMP.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR PRODOM; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
DR Homeobox; DNA-binding; Developmental protein; Nuclear protein.
FT DOMAIN 1 134 PRO-RICH.
FT DNA BIND 138 197 HOMEBOX.
SQ SEQUENCE 271 AA; 29986 MW; 376765557A71C962 CRC64;
Query Match 22.7%; Score 59; DB 1; Length 271;
Best Local Similarity 45.7%; Pred. No. 4.7;
Matches 16; Conservative 2; Mismatches 17; Indels 0; Gaps 0;
CY 1 BPLGRGDSPLRPPHLMQCGWRHSFSAPEL 35
DB 119 DPLGKPLMSPLRPLHKRKGQVRFSDQTVEL 153
RESULT 4
HMPH_HUMAN STANDARD; PRT; 270 AA.
AC Q03014;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Homeobox protein PRH (hematopoietically expressed homeobox) (Homeobox
DE protein HHEX).
GN HHEX OR PRHX OR PRH OR HEX.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=93087175; PubMed=1360645;

RA Crompton M.R., Bartlett T.J., Macgregor A.D., Manfioletti G.,
RA Buratti E., Giancotti V., Goodwin G.H.;
RT "Identification of a novel vertebrate homeobox gene expressed in
RT haematopoietic cells."
RL Nucleic Acids Res. 20:5661-5667(1992).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93384629; PubMed=8103988;
RA Hromas R.A., Collins S.J., Radich J.;
RT "PCR cloning of an orphan homeobox gene (PRH) preferentially
RT expressed in myeloid and liver cells."
RL Biochem. Biophys. Res. Commun. 195:976-983(1993).
[3]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE=22388257; PubMed=12477937;
RA Straube R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heile F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein W.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Boulford G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalley D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[4]
RP SEQUENCE OF 122-270 FROM N.A.
RX MEDLINE=93219088; PubMed=8096636;
RA Bedford F.K., Ashworth A., Enver T., Wiedemann L.M.;
RT "HEX: a novel homeobox gene expressed during haematopoiesis and
RT conserved between mouse and human."
RL Nucleic Acids Res. 21:1245-1249(1993).
[5]
RP 3D-STRUCTURE MODELING OF 136-196.
RX MEDLINE=9425190; PubMed=7911091;
RA Neidle S., Goodwin G.H.;
RT "A homology-based molecular model of the proline-rich homeodomain
RT protein PRH, from haematopoietic cells."
RL FEBS Lett. 345:93-98(1994).
CC -1- FUNCTION: RECOGNIZES THE DNA SEQUENCE 5'-ATTAA-3'. MAY PLAY A ROLE
CC IN HEMATOPOIETIC DIFFERENTIATION.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- TISSUE SPECIFICITY: LIVER AND PROMYELOCYTIC LEUKEMIA CELL
CC LINE HL60.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED DURING HEMATOPOIESIS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; X67235; CA47661.1; -.
DR EMBL; L16499; AAA02988.1; -.
DR EMBL; BC015110; AAH15110.1; -.
DR EMBL; Z21533; CA79730.1; -.
DR PIR; JN0767; JN0767.
DR HSSP; P22808; INK3.
DR TRANSFAC; T02092; -.
DR Genew; HGNC:4901; HHEX.

DR MIM: 604420; -
DR GO: 0003700; F:transcription factor activity; TMS.
DR GO: 0006960; P:antimicrobial humoral response (sensu inver. . .); TMS.
DR GO: 0007275; P:development; TMS.
DR InterPro: IPR001356; Homeobox.
DR Pfam: PF00046; homeobox.1.
DR PRINTS: PR00024; HOMEBOX.
DR PRODOM: P000010; Homeobox; 1.
DR SMART: SM00389; HOX; 1.
DR PROSITE: PS00027; HOMEBOX_1; 1.
DR PROSITE: PS00071; HOMEBOX_2; 1.
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein.
FT DOMAIN 1 133 PRO-RICH.
FT DNA_BIND 137 136 HOMEBOX.
FT CONFLICT 115 115 L -> V (IN REF. 2).
SQ SEQUENCE 270 AA; 30021 MW; 9C16B6D49475FC CRC64;

Query Match 22.3%; Score 58; DB 1; Length 270;
Best Local Similarity 50.0%; Pred. No. 6.1;
Matches 14; Conservative 2; Mismatches 12; Indels 0; Gaps 0;

Qy 1 EPLGPRGDSPLQRPQHLMDQGMKHS 28
Db 118 DPLGKPLMSPFLQRPPLHRKGGQVRFS 145

RESULT 5
NFC1 HUMAN STANDARD; PRT; 943 AA.
ID 095644; Q12865; Q15793;
AC 16-OCT-2001 (Rel. 40; Created)
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DT 15-SEP-2003 (Rel. 42; Last annotation update)
DE Nuclear factor of activated T-cells, cytoplasmic 1 (NFAT transcription
DE complex cytosolic component) (NF-ATc1) (NF-ATC).
GN NFATC1 OR NFATC OR NFAT2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
(1)
RP SEQUENCE FROM N.A. (ISOFORM A-ALPHA).
RC TISSUE=Peripheral blood lymphocytes, and T-cell;
RX MEDLINE=94261186; PubMed=8202141;
RA Northrop J.P., Ho S.N., Chen L., Thomas D.J., Timmerman L.A.,
RA Nolan G.P., Admon A., Crabtree G.R.;
RT "NF-AT components define a family of transcription factors targeted in
RT T-cell activation.";
RL Nature 369:497-502 (1994).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM B-BETA).
RC TISSUE=B-cell;
RX MEDLINE=96355439; PubMed=8702849;
RA Park J., Takeuchi A., Sharma S.;
RT "Characterization of a new isoform of the NFAT (nuclear factor of
RT activated T cells) gene family member NFATc.";
RL J. Biol. Chem. 271:20914-20921 (1996).
RN [3]
RN ERRATUM.
RA Park J., Takeuchi A., Sharma S.;
RL J. Biol. Chem. 271:33705-33705 (1996).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORMS A-ALPHA, B-ALPHA AND C-BETA).
RC TISSUE=B-cell lymphoma;
RX MEDLINE=99170294; PubMed=10072078;
RA Chuvpilo S., Zimmer M., Kerstan A., Gloeckner J., Avots A., Escher C.,
RA Fischer C., Inashkina I., Jankevics E., Berberich-Siebel F.,
RA Schmitt E., Serfling E.;
RT "Alternative polyadenylation events contribute to the induction of
RT NF-ATC in effector T cells.";
RL Immunity 10:261-269 (1999).
RN [5]
RP MUTAGENESIS.

RX MEDLINE=20119316; PubMed=10652349;
RA Porter C.M., Havens W.A., Clipstone N.A.;
RT "Identification of amino acid residues and protein kinases involved in
RT the regulation of NFATC subcellular localization.";
RL J. Biol. Chem. 275:3543-3551 (2000).
RN [6]
RP ALTERNATIVE SPLICING, AND CHARACTERIZATION.
RX MEDLINE=99288090; PubMed=10358178;
RA Chuvpilo S., Avots A., Berberich-Siebel F., Gloeckner J., Fischer C.,
RA Kerstan A., Escher C., Inashkina I., Hlubek F., Jankevics E.,
RA Brabler T., Serfling E.;
RT "Multiple NF-ATC isoforms with individual transcriptional properties
RT are synthesized in T lymphocytes.";
RL J. Immunol. 162:7294-7301 (1999).
RN [7]
RP REVIEW.
RX MEDLINE=99189746; PubMed=10089876;
RA Crabtree G.R.;
RT "Generic signals and specific outcomes: signaling through Ca2+,
RT calcineurin, and NF-AT.";
RL Cell 96:611-614 (1999).
CC -1- FUNCTION: Plays a role in the inducible expression of cytokine
CC genes in T cells, especially in the induction of the IL-2 or IL-4
CC gene transcription. Also control gene expression in embryonic
CC cardiac cells. Could regulate not only the activation and
CC proliferation but also the differentiation and programmed death of
CC T-lymphocytes as well as lymphoid and non-lymphoid cells.
CC -1- SUBUNIT: Member of the multicomponent NFATC transcription complex
CC that consists of at least two components, a pre-existing
CC cytoplasmic component NFATC2 and an inducible nuclear component
CC NFATC1. Other members such as NFATC4, NFATC3 or members of the
CC activating protein-1 family, MAP, GATA4 and Cbp/p300 can also bind
CC the complex. NFATC proteins bind to DNA as monomers.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic for the phosphorylated form and
CC nuclear after activation that is controlled by calcineurin-
CC mediated dephosphorylation. Rapid nuclear exit of NFATC is thought
CC to be one mechanism by which cells distinguish between sustained
CC and transient calcium signals. The subcellular localization of
CC NFATC play a key role in the gene transcription.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=6;
CC Comment=Isoform C-alpha and isoform C-beta are the strongest
CC activator of gene transcription, followed by isoform A-alpha and
CC isoform A-beta, whereas isoform B-alpha and isoform B-beta are
CC the weakest. Isoform B-alpha, isoform B-beta, isoform C-alpha
CC and isoform C-beta, both present in T-cells, can modulate their
CC transcriptional activity;
CC Name=C-alpha;
CC IsoId=O95644-1; Sequence=Displayed;
CC Note=An additional isoform may be produced by alternative
CC initiation at Met-37 of isoform C-alpha;
CC Name=A-alpha;
CC IsoId=O95644-2; Sequence=VSP_005591; VSP_005592;
CC Note=An additional isoform may be produced by alternative
CC initiation at Met-37 of isoform A-alpha;
CC Name=B-beta;
CC IsoId=O93644-3; Sequence=VSP_005590, VSP_005591, VSP_005592;
CC Name=B-alpha;
CC IsoId=O95644-4; Sequence=VSP_005593;
CC Note=An additional isoform may be produced by alternative
CC initiation at Met-37 of isoform B-alpha;
CC Name=B-beta;
CC IsoId=O93644-5; Sequence=VSP_005590, VSP_005593;
CC Name=C-beta;
CC IsoId=O95644-6; Sequence=VSP_005590;
CC Event=Alternative initiation;
CC Comment=6 isoforms may be produced by alternative initiation at
CC Met-1 and Met-37 of alpha-type isoforms;
CC TISSUE SPECIFICITY: Expressed in thymus, peripheral leukocytes as
CC T-cells and spleen. Isoforms A are preferentially expressed in
CC effector T-cells (thymus and peripheral leukocytes) whereas
CC isoforms B and isoforms C are preferentially expressed in naive T-
CC cells (spleen). Isoforms B are expressed in naive T-cells after

CC first antigen exposure and isoforms A are expressed in effector T-
 CC cells after second antigen exposure.
 CC -1- INDUCTION: Only isoforms A are inducibly expressed in T
 CC lymphocytes upon activation of the T-cell receptor (TCR) complex.
 CC Induced after co-addition of phorbol 12-myristate 13-acetate (PMA)
 CC and ionomycin. Also induced after co-addition of 12-O-
 CC tetradecanoylphorbol-13-acetate (TPA) and ionomycin. Weakly
 CC induced with PMA, ionomycin and cyclosporin A.
 CC -1- DOMAIN: Rel Similarity Domain (RSD) allows DNA-binding and
 CC cooperative interactions with API factors.
 CC -1- DOMAIN: The N-terminal transactivation domain (TAD-A) binds to and
 CC is activated by Cbp/p300. The dephosphorylated form contains two
 CC unmasked nuclear localization signals (NLS), which allow
 CC translocation of the protein to the nucleus.
 CC -1- DOMAIN: Isoforms C have a C-terminal part with an additional
 CC trans-activation domain, TAD-B, which acts as a transcriptional
 CC activator. Isoforms B have a shorter C-terminal part without
 CC complete TAD-B which acts as a transcriptional repressor.
 CC -1- PTM: Phosphorylated by NFATC-kinase; dephosphorylated by
 CC calcineurin.
 CC -1- SIMILARITY: BELONGS TO THE REL/DORSAL FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U08015; AAA19601.1; -
 CC EMBL: U59736; AAC50869.1; -
 CC EMBL: U80917; AAD00450.1; -
 CC EMBL: U80918; AAD00451.1; -
 CC EMBL: U80919; AAD00452.1; -
 CC PDB: 1A66; 17-JUN-98.
 CC PDB: 1NFA; 01-APR-97.
 CC TRANSFAC: T01945; -
 CC DR Genev: HGNC:7775; NFATC1.
 CC MIM: 600489; -
 CC GO: GO:0005737; C:cytoplasm; TAS.
 CC GO: GO:0005528; F:FK506 binding activity; TAS.
 CC GO: GO:0003700; F:transcription factor activity; TAS.
 CC GO: GO:0006366; P:transcription from Pol II promoter; TAS.
 CC InterPro: IPR002909; IPT.TIG.
 CC InterPro: IPR000451; NF_Rel_dor_fam.
 CC Pfam: PF00554; RHD; 1.
 CC Pfam: PF01833; TIG; 1.
 CC SMART: SM00429; IPT; 1.
 CC PROSITE: PS01204; REL_1; FALSE_NEG.
 CC PROSITE: PS50254; REL_2; 1.
 CC Transcription regulation; Activator; Repressor; Nuclear protein;
 CC DNA-binding; Alternative initiation; Alternative splicing;
 CC Phosphorylation; Repeat; 3D-structure.
 CC CHAIN 1 943
 CC NUCLEAR FACTOR OF ACTIVATED T-CELLS,
 CC CYTOPLASMIC 1.
 CC NUCLEAR FACTOR OF ACTIVATED T-CELLS,
 CC CYTOPLASMIC 1, ALPHA-TYPE ISOFORMS.
 CC FOR ALPHA-TYPE ISOFORMS.
 CC CALCINEURIN-BINDING.
 CC TRANS-ACTIVATION DOMAIN A (TAD-A).
 CC 3 X SP REPEATS.
 CC 1.
 CC 2.
 CC 3.
 CC NUCLEAR LOCALIZATION SIGNAL.
 CC NUCLEAR EXPORT SIGNAL.
 CC DNA-BINDING.
 CC NUCLEAR LOCALIZATION SIGNAL.
 CC NUCLEAR LOCALIZATION SIGNAL.
 CC TRANS-ACTIVATION DOMAIN B (TAD-B).
 CC NUCLEAR EXPORT SIGNAL.
 CC PHOSPHORYLATION.
 CC MPTSPVPVSKPPLPAAVFGRGELGAPAPAGGTMSAS
 FT VARSPLIC 1 42

FT E -> MTGLEDEDFEFLFEFNQDECAAAAP (in
 FT isoform A-beta, isoform B-beta and
 FT isoform C-beta).
 FT /Frid=VSP_005590.
 FT VPIIKTEPTDIDYEPATCG -> GNAIFLTYSREHERVGC
 FT F (in isoform A-alpha and isoform A-
 FT beta).
 FT /Frid=VSP_005591.
 FT Missing (in isoform A-alpha and isoform
 FT VARSPLIC 717 943
 FT Query Match 22.3%; Score 58; DB 1; Length 943;
 FT Best Local Similarity 34.5%; Pred. No. 26;
 FT Matches 19; Conservative 4; Mismatches 18; Indels 14; Gaps 2;
 Oy 2 PLPRGGQSPILQ-----RPOHMQGQMRHSFSAPELTL--RODKRP 42
 Db 850 PLPRATGPCTLQGCSPACPRATGPRQHLSTVRDSPTGPRLLPVEHEDGS 904
 RESULT 6
 ONC2_HUMAN STANDARD; PRT; 485 AA.
 AC O95948;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE One cut domain family member 2 (ONECUT-2 transcription factor) (OC-2).
 GN ONECUT2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A..
 RX MEDLINE=99115605; PubMed=9915796;
 RA Jacquemin P., Lannoy V., Rousseau G.G., Lemaigre F.P.;
 RT "OC-2, a novel mammalian member of the ONECUT class of homeodomain
 RT transcription factors whose function in liver partially overlaps with
 RT that of hepatocyte nuclear factor-6";
 RL J. Biol. Chem. 274:2665-2671(1999).
 CC -1- FUNCTION: TRANSCRIPTIONAL ACTIVATOR. ACTIVATES THE TRANSCRIPTION
 CC OF A NUMBER OF LIVER GENES SUCH AS HNF3B.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: Contains 1 CUT domain.
 CC -1- SIMILARITY: BELONGS TO THE CUT HOMEBOX FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Y18198; CAB38253.1; -
 CC TRANSFAC: T03259; -
 CC DR Genev: HGNC:8139; ONECUT2.
 CC MIM: 604894; -
 CC GO: GO:0003702; F:RNA polymerase II transcription factor acti. . .; TAS.
 CC GO: GO:0007397; P:histogenesis and organogenesis; TAS.
 CC InterPro: IPR007108; Cut homeo.
 CC InterPro: IPR003350; Homeo CUT.
 CC InterPro: IPR001356; Homeobox.
 CC Pfam: PF02376; CUT; 1.
 CC Pfam: PF00046; homeobox; 1.
 CC ProDom: PD000010; Homeobox; 1.
 CC SMART: SM00389; HOX; 1.
 CC PROSITE: PS00027; HOMEBOX 1; FALSE_NEG.
 CC PROSITE: PS50071; HOMEBOX 2; 1.
 CC Transcription regulation; Homeobox; DNA-binding; Nuclear protein;
 CC Activator.
 CC DNA_BIND 305 391 CUT.
 FT DNA_BIND 407 466 HOMEBOX.

```

FT DOMAIN 18 37 POLY-GLY.
FT DOMAIN 62 66 POLY-PRO.
FT DOMAIN 75 82 POLY-ALA.
FT DOMAIN 152 165 POLY-HIS.
FT DOMAIN 238 303 POLY-SER.
SQ SEQUENCE 485 AA; 52483 MW; AF21E052EFBE5DA1 CRC64;

Query Match
Best Local Similarity 36.4%; Score 57; DB 1; Length 485;
Matches 16; Conservative 4; Mismatches 24; Indels 0; Gaps 0;

Oy 6 RCSDSPLLQRPQHLMDQGMHSFSGAPPELLRQDKRPRSGSGTG 49
Db 260 RLGLPDPAMMGNHNGHHPGHTQSHGPVLPARSRERPRSSSGS 303

RESULT 7
ARS_PSEAE STANDARD; PRT; 535 AA.
ID P51691.
AC 01-OCT-1996 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Arylsulfatase (EC 3.1.6.1) (Aryl-sulfate sulphohydrolase).
GN ATSA OR PA0183.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OX NCBI_TaxId=287;
[1] SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RP STRAIN=ATCC 15692 / PA01;
RX MEDLINE=95262702, PubMed=7744061;
RA Beil S., Kehrli H., James P., Staudemann W., Cook A.M., Leisinger T.,
RA Kereesz M.A.;
RA "Purification and characterization of the arylsulfatase synthesized
RA by Pseudomonas aeruginosa PAO during growth in sulfate-free medium
RA and cloning of the arylsulfatase gene (atsa).";
RA Eur. J. Biochem. 229:385-394(1995).
[2]
[3] REVISIONS.
RA Kereesz M.A.;
RA Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reiter J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RA "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RA opportunistic pathogen.";
RA Nature 406:959-964 (2000).
[1] FUNCTION: SHOWS MAXIMAL ACTIVITY AT 57 DEGREES CELSIUS AND PH 8.9.
[2] INHIBITION THAT OVERSTEPS 20 MINUTES ABOVE 50 DEGREES CELSIUS
[3] LEADS TO ENZYME INACTIVATION.
CC -1- CATALYTIC ACTIVITY: A phenol sulfate + H(2)O = a phenol + sulfate.
CC -1- SUBUNIT: Monomer.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -1- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z48540; CAA08421.2; -

```

```

DR EMBL; AE004456; AAG03573.1; -
DR PIR; D83622; D83622.
DR PDB; 1HDH; 22-NOV-01.
DR InterPro; IPR000917; Sulfatase.
DR Pfam; PF00884; Sulfatase; 1.
DR PROSITE; PS00523; SULFATASE_1; 1.
DR PROSITE; PS00149; SULFATASE_2; 1.
KM Hydrolase; Complete proteome; 3d-structure.
FT INIT_MER 0
FT ACT_SITE 114 114 POTENTIAL.
FT CONFLICT 1 1 S->D (IN REF. 1; AA SEQUENCE).
SQ SEQUENCE 535 AA; 59814 MW; 7404F3749B14EA77 CRC64;

Query Match
Best Local Similarity 36.1%; Score 57; DB 1; Length 535;
Matches 13; Conservative 8; Mismatches 13; Indels 2; Gaps 1;

Oy 13 LQRPQHLMDQGMHSFSGAPPELLRQDKRPRSGSGTG 48
Db 213 LQAPREIVE--KYRGVDPAGPEALRQERLAKELG 246

RESULT 8
TYSY_HUMAN STANDARD; PRT; 312 AA.
ID P04818.
AC 13-AUG-1987 (Rel. 05, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Thymidylate synthase (EC 2.1.1.45) (TS) (TSase).
GN TYMS OR TS.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
[1] SEQUENCE FROM N.A.
RP MEDLINE=85215597; PubMed=2987839;
RX Takeishi K., Kanada S., Aynsawa D., Shimizu K., Gotoh O., Seno T.;
RT "Nucleotide sequence of a functional cDNA for human thymidylate
RT synthase.";
RL Nucleic Acids Res. 13:2035-2043(1985).
[2]
[3] SEQUENCE FROM N.A.
RX MEDLINE=91056070; PubMed=2243092;
RA Kanada S., Nalbantoglu J., Takeishi K., Shimizu K., Gotoh O.,
RA Seno T., Aynsawa D.;
RT "Structural and functional analysis of the human thymidylate synthase
RT gene.";
RL J. Biol. Chem. 265:20277-20284(1990).
[3]
RP SEQUENCE FROM N.A.
RX TISSUE=Bone marrow, and Placenta;
RC MEDLINE=22388257; PubMed=12477932;
RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Schenker C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhac N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusha K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedini T.B., Toshiyuki S., Carinini P., Prange C.,
RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunnarsson P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buterfield V.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";

```

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [4]
 RP SEQUENCE OF 1-67 FROM N.A.
 RX MEDLINE=90110051; PubMed=2532645;
 RA Takeishi K., Kaneda S., Ayusawa D., Shimizu K., Gotoh O., Seno T.;
 RT "Human thymidylate synthase gene: isolation of phage clones which
 RT cover a functionally active gene and structural analysis of the
 RT region upstream from the translation initiation codon.";
 RL J. Biochem. 106:575-583(1989).
 RN [5]
 RP SEQUENCE OF 1-24.
 RX MEDLINE=85261174; PubMed=3839505;
 RA Shimizu K., Ayusawa D., Takeishi K., Seno T.;
 RT "Purification and NH2-terminal amino acid sequence of human
 RT thymidylate synthase in an overproducing transformant of mouse FM3A
 RT cells.";
 RL J. Biochem. 97:845-850(1985).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
 RX MEDLINE=96110704; PubMed=8845352;
 RA Schiffer C.A., Clifton I.J., Davieson V.J., Santi D.V., Stroud R.M.;
 RT "Crystal structure of human thymidylate synthase: a structural
 RT mechanism for guiding substrates into the active site.";
 RL Biochemistry 34:16279-16287(1995).
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
 RX MEDLINE=21229106; PubMed=11329255;
 RA Phan J., Koli S., Minor W., Dunlap R.B., Berger S.H., Leblond L.;
 RT "Human thymidylate synthase is in the closed conformation when
 RT complexed with dUMP and raltitrexed, an antifolate drug.";
 RL Biochemistry 40:1897-1902(2001).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
 RX MEDLINE=21216721; PubMed=11278511;
 RA Phan J., Steadman D.J., Koli S., Ding W.C., Minor W., Dunlap R.B.,
 RA Berger S.H., Leblond L.;
 RT "Structure of human thymidylate synthase suggests advantages of
 RT chemotherapy with noncompetitive inhibitors.";
 RL J. Biol. Chem. 276:14170-14177(2001).
 CC -I- CATALYTIC ACTIVITY: 5,10-methylenetetrahydrofolate + dUMP =
 CC dihydrofolate + dTMP.
 CC -I- PATHWAY: Deoxyribonucleotide biosynthesis.
 CC -I- SUBUNIT: Homodimer.
 CC -I- SIMILARITY: BELONGS TO THE THYMIDYLATE SYNTHASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X02308; CAA26178.1; -;
 DR EMBL; D00596; BAA00472.1; -;
 DR EMBL; BC002567; AA02567.1; -;
 DR EMBL; BC013919; AA013919.1; -;
 DR EMBL; D00517; BAA00404.1; -;
 DR PIR; A23047; YXHUT.
 DR PDB; 1HWY; 31-JAN-01.
 DR PDB; 1HW3; 09-MAY-01.
 DR PDB; 1HW4; 09-MAY-01.
 DR PDB; 1HW2; 09-MAY-01.
 DR PDB; 1T00; 09-MAY-01.
 DR PDB; 1T06; 14-NOV-01.
 DR PDB; 1T0U; 19-SEP-01.
 DR GeneW; HGNC:12441; TYMS.
 DR MIM; 188350; -;
 DR GO; GO:0009157; P:deoxyribonucleoside monophosphate biosynthesis; TAS.
 DR GO; GO:006113; P:nucleobase, nucleoside, nucleotide and nucl. . .; TAS.
 DR InterPro; IPR000398; Thymidylat_synct.
 DR Pfam; PF00303; thymidylat_synct.1.
 DR PRINTS; PR00108; THYMSNTASE.

DR ProDom; PD001180; Thymidylat_synct. 1.
 DR PROSITE; PS00091; THYMIDYLATE SYNTHASE; 1.
 KW Transferase; Methyltransferase; Nucleotide biosynthesis; 3D-structure.
 FT INT MET 0 0
 FT ACT SITE 194 194
 FT TURN 28 28
 FT TURN 29 42
 FT HELIX 44 46
 FT STRAND 49 50
 FT TURN 54 65
 FT TURN 67 68
 FT TURN 74 77
 FT TURN 80 91
 FT TURN 92 93
 FT STRAND 96 96
 FT HELIX 97 101
 FT TURN 102 104
 FT TURN 107 109
 FT HELIX 110 112
 FT HELIX 114 119
 FT TURN 120 121
 FT TURN 123 124
 FT TURN 127 128
 FT STRAND 129 129
 FT HELIX 134 140
 FT STRAND 141 141
 FT STRAND 142 142
 FT TURN 148 149
 FT TURN 153 154
 FT STRAND 157 157
 FT HELIX 159 169
 FT TURN 171 172
 FT TURN 174 175
 FT STRAND 177 179
 FT TURN 183 185
 FT HELIX 186 188
 FT STRAND 195 203
 FT TURN 204 205
 FT STRAND 206 217
 FT TURN 218 220
 FT TURN 221 231
 FT TURN 240 241
 FT STRAND 243 257
 FT HELIX 258 260
 FT HELIX 261 268
 FT TURN 269 269
 FT STRAND 277 280
 FT HELIX 287 289
 FT HELIX 292 294
 FT STRAND 295 298
 SQ SEQUENCE 312 AA; 35584 MW; A66F0E6D1973AB41 CRC64;
 Query Match 21.7%; Score 56.5; DB 1; Length 312;
 Best Local Similarity 34.0%; Pred. No. 11;
 Matches 16; Conservative 7; Mismatches 19; Indels 5; Gaps 2;
 QY 2 PLGPRGDSPLRGPRH--LMDGOMRHSFAGSELLRQDKRRPRSGS 46
 DB 11 PLPPAQRDAEP RPFGELQYLQIQHILRCG--VRKDRGTGTGT 54
 RESULT 9
 ID SP7_HUMAN STANDARD; PRT; 431 AA.
 AC Q8TDD2;
 DT 15-SEP-2003 (Rel. 42, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DE Transcription factor Sp7 (Zinc finger protein osterix).
 GN SP7 OR OSX.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

OK NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Osteosarcoma;
RA Ganss B.W.;
RT "CDNA sequence, gene structure and chromosomal localization of the
   human osteix (OSX) gene";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Nakeshima K., Zhou X., de Crombrughe B.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Transcriptional activator essential for osteoblast
   differentiation. Binds to Sp1 and EKLF consensus sequences and to
   other G/C-rich sequences (by similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: BELONGS TO THE SP1 FAMILY OF C2H2-TYPE ZINC-FINGER
   PROTEINS.
CC -1- SIMILARITY: Contains 3 C2H2-type zinc fingers.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
   between the Swiss Institute of Bioinformatics and the EMBL outstation -
   the European Bioinformatics Institute. There are no restrictions on its
   use by non-profit institutions as long as its content is in no way
   modified and this statement is not removed. Usage by and for commercial
   entities requires a license agreement (See http://www.isb-sib.ch/announce/
   or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF477981; AAL84281.1; -
DR EMBL: AF466179; AAL33377.1; -
DR Genew: HGNC:17321; SP7.
DR MIM: 606633; -
DR InterPro: IPR007087; Znf_C2H2.
DR Pfam: PF00096; Zf-C2H2; 3.
DR ProDom: PDO00003; Znf_C2H2; 1.
DR SMART: SMO0355; Znf_C2H2; 3.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 3.
DR PROSITE: PS50157; ZINC_FINGER_C2H2_2; 3.
DR TRANSCRIPT: PS00028; ZINC_FINGER_C2H2_2; 3.
DR KW Nuclear protein; Repeat; Activator.
DR ZN FING 294 318 C2H2-TYPE 1.
DR ZN FING 324 348 C2H2-TYPE 2.
DR ZN FING 354 376 C2H2-TYPE 3.
DR FT 431 AA; 44994 MW; 454A6FEA84309FF9 CRC64;
SQ SEQUENCE
Query Match 21.7%; Score 56.5; DB 1; Length 431;
Best Local Similarity 35.0%; Pred. No. 16;
Matches 21; Conservative 4; Mismatches 14; Indels 21; Gaps 4;
QY 2 PLGPRGQDSP-LLO-RPQHLMQ-----GQMRHSFSAQFELRLQDKRPPSGSGTG 48
DB 204 PLNPAPYPAHLLQPGQHVLPQDYVKPRAVGNSQLGSSGAKP-----PGASTG 255

```

```

RT mitosis.";
RL Nature 380:544-547 (1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA MEDLINE=22388257; PubMed=12477932;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,
   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
   Diatchenko L., Maruina K., Farmer A.A., Rubin G.M., Hong L.,
   Stapleton M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
   Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
   Bosack S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
   Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
   Fahey J., Hellon E., Kettleman M., Madan A., Rodriguez S., Sanchez A.,
   Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,
   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
   Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
   Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
   "Generation and initial analysis of more than 15,000 full-length
   human and mouse cDNA sequences.";
   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (1.35 ANGSTROMS).
RX MEDLINE=97344079; PubMed=9200606;
RA Ranganathan R., Lu K.P., Hunter T., Noel J.P.;
RT "Structural and functional analysis of the mitotic rotamase Pim1
   suggests substrate recognition is phosphorylation dependent.";
   Cell 89:875-886 (1997).
CC -1- FUNCTION: Essential PPIase that regulates mitosis presumably by
   interacting with NIMA and attenuating its mitosis-promoting
   activity. Displays a preference for an acidic residue N-terminal
   to the isomerized proline bond. Catalyzing pSer/thr-Pro cis/trans
   isomerizations.
CC -1- CATALYTIC ACTIVITY: Peptidylproline (omega=180) = peptidylproline
   (omega=0).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: BELONGS TO THE PPIC/PARYULIN FAMILY OF ROTAMASES.
CC -1- SIMILARITY: Contains 1 WW domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
   between the Swiss Institute of Bioinformatics and the EMBL outstation -
   the European Bioinformatics Institute. There are no restrictions on its
   use by non-profit institutions as long as its content is in no way
   modified and this statement is not removed. Usage by and for commercial
   entities requires a license agreement (See http://www.isb-sib.ch/announce/
   or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U49070; AAC50492.1; -
DR EMBL: BC002899; AAH02899.1; -
DR PIR: S68520; S68520.
DR PDB: 1PIN; 25-NOV-98.
DR PDB: 1F8A; 23-AUG-00.
DR PDB: 1I6C; 18-JUL-01.
DR PDB: 1I8G; 18-JUL-01.
DR PDB: 1I8H; 18-JUL-01.
DR Genew: HGNC:8988; PIN1.
DR GK: Q1526; -
DR MIM: 601052; -
DR GO: GO:0005634; C:nucleus; TAS.
DR GO: GO:0005515; F:protein binding activity; TAS.
DR GO: GO:0007088; P:regulation of mitosis; TAS.
DR InterPro: IPR000297; Rotamase.
DR InterPro: IPR001202; WW_Rsp5_WWP.
DR Pfam: PF00639; Rotamase; 1.
DR Pfam: PF00397; WW; 1.
DR SMART: SMO0456; WW; 1.
DR PROSITE: PS01096; PPIC_PPIASE_1; 1.
DR PROSITE: PS50198; PPIC_PPIASE_2; 1.

```

DR PROSITE; PS01159; MW DOMAIN 1; 1.
 DR PROSITE; PS50020; MW DOMAIN 2; 1.
 KW Isomerase; Rotamase; Nuclear protein; Cell cycle; 3D-structure.
 FT DOMAIN 5 39 MW.
 FT TURIN 52 163 PPIC.
 FT TURIN 9 10
 FT TURIN 11 15
 FT TURIN 17 18
 FT STRAND 22 26
 FT TURIN 27 29
 FT STRAND 32 33
 FT STRAND 55 62
 FT TURIN 65 66
 FT STRAND 72 72
 FT TURIN 73 74
 FT STRAND 75 75
 FT HELIX 82 98
 FT TURIN 99 99
 FT HELIX 103 110
 FT HELIX 114 118
 FT TURIN 119 120
 FT STRAND 121 125
 FT TURIN 127 128
 FT HELIX 132 140
 FT TURIN 143 144
 FT STRAND 146 146
 FT STRAND 150 151
 FT STRAND 156 161
 SO SEQUENCE 163 AA; 18243 MW; 35391AF40B7D1E13 CRC64;
 Query Match 21.5%; Score 56; DB 1; Length 163;
 Best Local Similarity 36.6%; Pred. No. 6;
 Matches 15; Conservative 5; Mismatches 17; Indels 4; Gaps 1;
 Qy 4 GPRGDSPLQRPQHLMDQGMKHSFGPELLRQDKPKRS 44
 Db 45 GKNGGEPARVRCSHLT---VKHSQSRPSPSWRQEKITRT 81
 RESULT 11
 HMPH_CHICK STANDARD; PRT; 277 AA.
 ID HMPH_CHICK
 AC Q05502;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Homeobox protein PRH.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93087175; PubMed=1360645;
 RA Burattini E., Giacchetti V., Goodwin G.H.;
 RT "Identification of a novel vertebrate homeobox gene expressed in
 RT hematopoietic cells.";
 RL Nucleic Acids Res. 20:5661-5667(1992).
 CC -1- FUNCTION: RECOGNIZES THE DNA SEQUENCE 5'-ATTAA-3'. MAY PLAY A ROLE
 CC IN HEMATOPOIETIC DIFFERENTIATION.
 CC IN HEMATOPOIETIC DIFFERENTIATION.
 CC -1- TISSUE SPECIFICITY: NUCLEAR (Probable).
 CC PERIPHERAL BLOOD ERYTHROCYTES AND IN THE LIVER AND LUNG.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; X64711; CAA45966.1; -.
 DR PIR; S78063; S78063.
 DR HSSP; P22808; INK3.
 DR TRANSFAC; T02091; -.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox.1.
 DR PRINTS; PR00024; HOMEBOX.
 DR ProDom; PD000010; Homeobox.1.
 DR SMART; SM00389; HOX.1.
 DR PROSITE; PS00027; HOMEBOX 1; 1.
 DR PROSITE; PS50071; HOMEBOX 2; 1.
 KW Homeobox; DNA-binding; Developmental protein; Nuclear protein.
 FT DOMAIN 1 140
 FT DNA BIND 144 203 HOMEBOX.
 SO SEQUENCE 277 AA; 30213 MW; BE744C143FB9F9FC CRC64;
 Query Match 21.5%; Score 56; DB 1; Length 277;
 Best Local Similarity 46.4%; Pred. No. 11;
 Matches 13; Conservative 3; Mismatches 12; Indels 0; Gaps 0;
 Qy 1 EPLGPRGDSPLQRPQHLMDQGMKHS 28
 Db 125 DPLGKPLMSPFIQPLPKRKGQVRS 152
 RESULT 12
 MENE_SALTY STANDARD; PRT; 455 AA.
 ID MENE_SALTY
 AC P37418;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE O-succinylbenzoic acid-CoA ligase (EC 6.2.1.26) (OSB-CoA synthetase)
 DE (O-succinylbenzoate-CoA synthase).
 GN MENE OR STM2305.
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxID=602;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=L72 / SGSC1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609;
 RA McLelland M., Sanderson K.E., Spiegh J., Clifton S.W., Latreille P.,
 RA Courtney L., Porwollik S., Ali D., Dance M., Du P., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
 RT L72.";
 RL Nature 413:852-856(2001).
 RN [2]
 RP SEQUENCE OF 228-455 FROM N.A.
 RC STRAIN=L72;
 RX MEDLINE=94266712; PubMed=8206837;
 RA Roland K.L., Escher C.R., Spitznagel J.K.;
 RT "Isolation and characterization of a gene, pmrD, from Salmonella
 RT typhimurium that confers resistance to polymyxin when expressed in
 RT multiple copies.";
 RL J. Bacteriol. 176:3589-3597(1994).
 CC -1- FUNCTION: O-SUCCINYLBENZOIC ACID (OSB) TO O-SUCCINYLBENZOYL-COA
 CC (OSB-COA).
 CC -1- CATALYTIC ACTIVITY: ATP + O-succinylbenzoate + CoA = AMP +
 CC diphosphate + O-succinylbenzoyl-CoA.
 CC -1- PATHWAY: Menadiquinone biosynthesis.
 CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL, AE008803; AAA21206.1; -
 CC EMBL, U02281; AAA21323.1; -
 CC StyGene; SGI0221; meme.
 DR InterPro; IPR000873; AMP-bind.
 DR Pfam; PF00501; AMP-binding; 1.
 DR PROSITE; PS00455; AMP BINDING; 1.
 KW Menaglutone biosynthesis; Ligase; Complete proteome.
 SQ SEQUENCE 455 AA; 50244 MW; BC830AB7BFAB0F42 CRC64;

Query Match 21.3%; Score 55.5; DB 1; Length 455;
 Best Local Similarity 40.5%; Pred. No. 23;
 Matches 17; Conservative 4; Mismatches 18; Indels 3; Gaps 2;

2 PLGPRGDSPLRPO-HLMDGOMRHSFASAPPELLRQKRP 42
 171 PGGP--QDDWLSLPLFHVSGGIMRWLFAQAKMTVRDKOP 210

RESULT 13
 ID ABS_HUMAN STANDARD; PRT; 622 AA.
 AC Q9UTV9; Q96BK5; Q96K05; Q9NM04;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE DEAD-box protein abstract homolog.
 GN ABS.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OK NCBI_TaxID=9606;
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20076660; PubMed=10607561;
 RA "Developmental and cell biological functions of the Drosophila DEAD-box protein abstract."
 RT Curr. Biol. 9:1373-1381(1999).
 RL [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
 RA Wagaatsuma M., Hosokiri T., Kaku Y., Kodaira H., Kondo Y., Sugawara M.,
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
 RA Matenabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saio K.,
 RA Yamamoto U., Wakamatsu A., Nakamura Y., Nagahara K., Masuno Y.,
 RA Ninomiya K., Iwayanagi T.;
 RT "NEO human cDNA sequencing project."
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=skin;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo W.F., Caesavant T.L., Scheetz T.E.,
 RA Brownstein M.U., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullighy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Paaby J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Small D.E.,
 RA Scherch A., Schain J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -!- FUNCTION: PROBABLE ATP-DEPENDENT RNA HELICASE. IS REQUIRED DURING
 CC POST-TRANSCRIPTIONAL GENE EXPRESSION.
 CC -!- SUBCELLULAR LOCATION: Nuclear (potential).
 CC -!- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL, AF195417; AAF04150.1; -
 DR EMBL, AK001255; BAA91585.1; -
 DR EMBL, AK027768; BAB55355.1; -
 DR EMBL, BC015476; AAH15476.1; -
 DR HSSP; Q58083; 1HV8.
 DR GO; GO:0006915; P:apoptosis; TAS.
 DR GO; GO:0007275; P:development; TAS.
 DR GO; GO:0006396; P:RNA processing; TAS.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR000629; DEAD_box.
 DR InterPro; IPR001650; Helicase_C.
 DR InterPro; IPR001878; znf_CCHC.
 DR Pfam; PF00270; DEAD; 1.
 DR Pfam; PF00271; Helicase_C; 1.
 DR Pfam; PF00098; zf-CCHC; 1.
 DR SMART; SM00487; DEXDC; 1.
 DR SMART; SM00490; HELIC_C; 1.
 DR SMART; SM00343; ZNF_C2HC; 1.
 DR PROSITE; PS00039; DEAD ATP HELICASE; FALSE NEG.
 KW ATP-binding; RNA-binding; Helicase; Nuclear protein.
 FT NP_BIND 225 232
 FT SITE 344 347
 FT CONFLICT 17 43
 FT FT DEAD BOX.
 FT FT PAGGRSRSEADEDEDDVYVPLRQR -> LPERRARRK
 FT FT MTRTRTCPCMCVAP (IN REF. 1).
 FT FT CONFLICT 56 56 K -> E (IN REF. 2; BAA91585).
 FT FT CONFLICT 64 64 Q -> E (IN REF. 2; BAB55355).
 FT FT CONFLICT 165 165 K -> E (IN REF. 2; BAB55355).
 FT FT CONFLICT 191 191 A -> T (IN REF. 2; BAA91585).
 FT FT CONFLICT 352 352 M -> T (IN REF. 2; BAA91585).
 FT FT CONFLICT 552 552 L -> Q (IN REF. 2; BAB55355).
 SQ SEQUENCE 622 AA; 69837 MW; EDA328724E0DF99A CRC64;

Query Match 21.3%; Score 55.5; DB 1; Length 622;
 Best Local Similarity 31.9%; Pred. No. 33;
 Matches 15; Conservative 10; Mismatches 21; Indels 1; Gaps 1;

2 PLGPRGDSPLRPO-HLMDGOMRHSFASAPPELLRQKRP 48
 78 PLGPGSNVS-LLDDHQLKERAKEAKESAKELKEKILLESVAEG 123

RESULT 14
 ID RL18 YEAST STANDARD; PRT; 186 AA.
 AC P07279;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE 60S ribosomal protein l18 (RP28).
 GN (RP118A OR RP28A OR YOL120C) AND (RP118B OR RP28B OR YNL301C OR
 GN N0425).
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetaceae; Saccharomycetaceae; Saccharomycetes.


```
DR Pfam: PF00090; TSP_1; 11.  
DR SMART; SM00209; TSP1; 12.  
DR PROSITE; PSS0215; ADAM_MEPRO; 1.  
DR PROSITE; PSS0546; CYSTEINE_SWITCH; FALSE NEG.  
DR PROSITE; PSS0427; DISINTEGRIN_1; FALSE NEG.  
DR PROSITE; PSS0214; DISINTEGRIN_2; FALSE NEG.  
DR PROSITE; PSS0092; TSP1; 12.  
DR PROSITE; PSS0142; ZINC_PROTEASE; 1.  
KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;  
KW Repeat; Extracellular matrix; Alternative splicing.  
FT SIGNAL; 1 18  
FT PROPEP 1 287  
FT CHAIN 288 1629  
FT DOMAIN 293 499  
FT DOMAIN 509 587  
FT DOMAIN 588 643  
FT DOMAIN 878 936  
FT DOMAIN 645 752  
FT DOMAIN 753 880  
FT DOMAIN 997 1054  
FT DOMAIN 1055 1104  
FT DOMAIN 1107 1166  
FT DOMAIN 1182 1240  
FT DOMAIN 1241 1296  
FT DOMAIN 1328 1379  
FT DOMAIN 1382 1440  
FT DOMAIN 1441 1494  
FT DOMAIN 1497 1555  
FT DOMAIN 1556 1613  
FT DOMAIN 88 96  
FT SITE 223 434  
FT METAL 434 434  
FT ACT_SITE 435 435  
FT METAL 438 438  
FT METAL 444 444  
FT CARBOHYD 112 112  
FT CARBOHYD 135 135  
FT CARBOHYD 271 271  
FT CARBOHYD 749 749  
FT CARBOHYD 840 840  
FT CARBOHYD 1213 1213  
FT CARBOHYD 1267 1267  
FT VARSPLIC 1064 1072  
FT  
FT VARSPLIC 1073 1629  
FT  
FT CONFLICT 367 367  
FT SEQUENCE 1629 AA; 182649 MW; C1C4CEPF588941F CRC64;  
SQ  
Query Match 21.2%; Score 55; DB 1; Length 1629;  
Best Local Similarity 35.9%; Pred. No. 1.1e+02;  
Matches 14; Conservative 3; Mismatches 16; Indels 6; Gaps 1;  
OY 8 QDSPLQRQHMLDQGMHSFAGPELLRQDKRPRSGS 46  
DB 1288 QDCSWSPCQRTPTDGLAQHPQ-----NEDYRPSAS 1320  
Search completed: July 24, 2003, 11:54:28  
Job time : 4.22401 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 11:47:53 ; Search time 10.9048 Seconds
(without alignments)
1159.539 Million cell updates/sec

Title: US-09-991-681-31
Perfect score: 260
Sequence: 1 EPLGPRGDSPLQRPQHLM.....SAGPELLRQDKRPRSGSTGS 49

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*
15: SP viirus:*
16: SP_bacteriap:*
17: SP_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	260	100.0	589	4	Q8N4Y4	Q8N4Y4 homo sapien
2	260	100.0	592	4	Q96CH9	Q96CH9 homo sapien
3	260	100.0	1770	4	Q9ULH6	Q9ULH6 homo sapien
4	63	22.2	462	16	Q9HUI9	Q9HUI9 pseudomonas
5	62.5	22.0	1655	5	Q24754	Q24754 drosophila
6	61	23.5	384	16	Q8DL53	Q8DL53 streptococc
7	60	23.1	293	2	Q9X6T3	Q9X6T3 streptococ
8	60	22.1	333	16	Q9RTI1	Q9RTI1 deinococcus
9	59.5	22.9	345	17	Q9YBM1	Q9YBM1 aeropyrum p
10	59	22.7	271	11	Q9RI12	Q9RI12 mus musculu
11	59	22.7	280	11	Q9CRV1	Q9CRV1 mus musculu
12	59	22.7	381	16	Q9ACT4	Q9ACT4 streptococ
13	59	22.7	644	4	Q9NR66	Q9NR66 homo sapien
14	58.5	22.5	276	12	Q85275	Q85275 potato viru
15	58.5	22.5	580	3	Q9C284	Q9C284 neurospora
16	58.5	22.5	778	3	Q96WJ2	Q96WJ2 kluyveromyc

17	58.5	22.5	3419	5	Q9W425	Q9W425 drosophila
18	58.5	22.5	3429	5	Q24593	Q24593 drosophila
19	58	22.3	172	11	Q9CUI2	Q9CUI2 mus musculu
20	58	22.3	221	6	Q95X08	Q95X08 sus scrofa
21	58	22.3	270	4	Q96C89	Q96C89 homo sapien
22	58	22.3	357	16	Q9A875	Q9A875 caulobacter
23	58	22.3	523	5	Q9N310	Q9N310 caenorhabdi
24	58	22.3	628	12	Q56253	Q56253 turnip yell
25	58	22.3	726	5	Q8MX55	Q8MX55 caenorhabdi
26	58	22.3	743	5	Q9N313	Q9N313 caenorhabdi
27	57.5	22.1	405	16	Q8UB85	Q8UB85 agrobacteri
28	57.5	22.1	500	13	Q8UWQ9	Q8UWQ9 gallus gall
29	57.5	22.1	511	16	Q8UWQ9	Q8UWQ9 xanthomonas
30	57.5	22.1	1207	5	Q8I174	Q8I174 drosophila
31	57.5	22.1	1451	10	Q49143	Q49143 arabidopsis
32	57.5	22.1	1451	10	Q49140	Q49140 arabidopsis
33	57.5	22.1	1451	10	Q49142	Q49142 arabidopsis
34	57.5	22.1	1451	10	Q49142	Q49142 arabidopsis
35	57.5	22.1	1675	13	Q8UUR1	Q8UUR1 gallus gall
36	57	21.9	159	13	Q8BYT7	Q8BYT7 mus musculu
37	56.5	21.7	119	11	Q9CTL8	Q9CTL8 xenopus lae
38	56.5	21.7	165	11	Q9CUI9	Q9CUI9 mus musculu
39	56.5	21.7	230	4	Q8WYK4	Q8WYK4 homo sapien
40	56.5	21.7	279	4	Q8WYK3	Q8WYK3 homo sapien
41	56.5	21.7	431	4	Q8TDD2	Q8TDD2 homo sapien
42	56.5	21.7	480	13	Q91046	Q91046 gobius sp.
43	56.5	21.7	561	16	Q8G4W0	Q8G4W0 bifidobacte
44	56.5	21.7	714	3	Q8WZX7	Q8WZX7 neurospora
45	56.5	21.7	807	10	Q9X107	Q9X107 arabidopsis

ALIGNMENTS

RESULT 1
ID Q8N4Y4 PRELIMINARY; PRT; 589 AA.
AC Q8N4Y4;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Similar to KIAA1244 protein (fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strausberg R.;
DR EMBL; BC033191; AAA33191.1; -.
FT NON_TER
SQ SEQUENCE 589 AA; 66086 MW; 8041BEA348DB65F7 CRC64;

Query Match 100.0%; Score 260; DB 4; Length 589;
Best Local Similarity 100.0%; Pred. No. 6.5e-26;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 EPLGPRGDSPLQRPQHLMQGMHRSFSGPELLRQDKRPRSGSTGS 49
Db 464 EPLGPRGDSPLQRPQHLMQGMHRSFSGPELLRQDKRPRSGSTGS 512
RESULT 2
AC Q96CH9 PRELIMINARY; PRT; 592 AA.
ID Q96CH9;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Struhsberg R.;
RT Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC014227; AAH14227.1; -.
KW Hypothetical protein.
FT NON TER
SQ SEQUENCE 592 AA; 66400 MW; F4A1E807B0DF47B5 CRC64;

Query Match
Best Local Similarity 100.0%; Score 260; DB 4; Length 592;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPLGRGQDSPILQRPHLMDOGQMRHSFSAPELLRQKRRSGSTGS 49
DB 467 EPLGRGQDSPILQRPHLMDOGQMRHSFSAPELLRQKRRSGSTGS 515

RESULT 3
ID Q9ULH6 PRELIMINARY; PRT; 1770 AA.
AC Q9ULH6; Q96PA6;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE BIG3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Hong W.;
RT "K1A144 as a novel distantly related member (BIG3) of the BIG1/Sec7p
RL Subfamily of ARF GEFs";
DR EMBL; AF411080; AAL04174.1; -.
DR InterPro; IPR000904; Sec7.
DR SMART; SM00222; Sec7; 1.
KW Hypothetical protein.
SQ SEQUENCE 1770 AA; 195845 MW; SE996E36A6F92AB4 CRC64;

Query Match
Best Local Similarity 100.0%; Score 260; DB 4; Length 1770;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPLGRGQDSPILQRPHLMDOGQMRHSFSAPELLRQKRRSGSTGS 49
DB 1645 EPLGRGQDSPILQRPHLMDOGQMRHSFSAPELLRQKRRSGSTGS 1693

RESULT 4
ID Q9HUI9 PRELIMINARY; PRT; 462 AA.
AC Q9HUI9;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Probable two-component response regulator.
GN PAS166.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,

```

```

RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrock-Adman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Latbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RT Nature 406:959-964(2000).
CC -1- SIMILARITY: THE N-TERMINAL REGION IS SIMILAR TO THAT OF OTHER
CC REGULATORY COMPONENTS OF SENSORY TRANSDUCTION SYSTEMS.
CC -1- SIMILARITY: CONTRAINS A SIGMA-54 FACTOR INTERACTION ATP-BINDING
CC DOMAIN.
CC EMBL; AE004929; AAG08551.1; -.
DR HSPF; P10958; IDPM.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR002197; HTH_Fib.
DR InterPro; IPR001789; Response_reg.
DR InterPro; IPR002078; Sig54_interact.
DR Pfam; PF02954; HTH_8; 1.
DR Pfam; PF00072; response_reg; 1.
DR Pfam; PF00158; Sigma54_activat; 1.
DR PRINTS; PR01590; HTHFIS.
DR ProDom; PD000039; Response_reg; 1.
DR SMART; SM00382; AAA; 1.
DR SMART; SM00448; REC; 1.
DR TIGRfam; TIGR01199; HTH_fib; 1.
DR PROSITE; PS50110; RESPONSE_REGULATORY; 1.
DR PROSITE; PS00675; SIGMA54_INTERACT_1; 1.
DR PROSITE; PS00676; SIGMA54_INTERACT_2; 1.
DR PROSITE; PS00688; SIGMA54_INTERACT_3; 1.
DR PROSITE; PS50045; SIGMA54_INTERACT_4; 1.
KW ATP-binding; DNA-binding; Phosphorylation; Sensory transduction;
KW Transcription regulation; Complete proteome.
KW SEQUENCE 462 AA; 50864 MW; 981D88847B2BABF CRC64;

Query Match
Best Local Similarity 48.3%; Score 63; DB 16; Length 462;
Matches 14; Conservative 3; Mismatches 10; Indels 2; Gaps 1;

QY 2 PLGRGQDSPILQRPHLMDOGQMRHSFSAPELLRQKRRSGSTGS 30
DB 322 PLRGERGDIPLFR--HFAEAGAMRHGLT 348

RESULT 5
ID Q24754 PRELIMINARY; PRT; 1655 AA.
AC Q24754;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Maetermind.
GN MAM.
OS Drosophila virilis (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7244;
RN [1]
RP SEQUENCE OF 260-762 FROM N.A.
RX MEDLINE=91251140; PubMed=1904096;
RA Newfeld S.J., Smollier D.A., Yedvobnick B.;
RT "Interpectic comparison of the unusually repetitive Drosophila locus
RT maetermind.";
RT J. Mol. Evol. 32:415-420(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94111143; PubMed=8283480;
RA Newfeld S.J., Schmid A.T., Yedvobnick B.;
RT "Homopolymer length variation in the Drosophila gene maetermind.";
RN [3]
RP SEQUENCE FROM N.A.

```

RC MEDLINE=94365848; PubMed=8083889;
 RA Newfield S.J., Tachida H., Yedvobnick B.;
 RT "Drive-selection equilibrium: homopolymer evolution in the *Drosophila*
 RT gene mastermind";
 RL J. Mol. Evol. 38:637-644 (1994).
 RN (4)
 RP SEQUENCE FROM N.A.
 RA Newfield S.J.;
 RL Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.
 DR EMBL; M92914; AAC37201.1; -;
 DR FlyBase; FBgn0011119; Dc11mam.
 SO SEQUENCE 1655 AA; 175048 MW; 247D8F8DF0DDDE12 CRC64;

Query Match 24.0%; Score 62.5; DB 5; Length 1655;
 Best Local Similarity 27.6%; Pred. No. 27;
 Matches 16; Conservative 9; Mismatches 22; Indels 11; Gaps 1;

OY 2 PLRGQDSPLLORPHOIMDOGQMRHSFSAG-----PELLRQDKRPRSGSTG 48
 Db 1451 PNGPQMQLTPAQMGOQHRRQQQQQHMGGGGGGGGMQMQLLQQQNAAGGGG 1508

RESULT 6
 Q8DL53 PRELIMINARY; PRT; 384 AA.
 AC Q8DL53;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE DevC-like ABC transporter permease protein.
 GN TLL0644.
 OS *Synechococcus elongatus* (Thermosynechococcus elongatus).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
 OX NCBI_TaxID=32046;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=BP-1;
 RX MEDLINE=22225144; PubMed=12240834;
 RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
 RA Watanabe A., Iriyuchi M., Kawashima K., Kimura T., Kishida Y.,
 RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
 RA Shimo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the thermophilic cyanobacterium
 RT *Thermosynechococcus elongatus* BP-1";
 RL DNA Rep. 9:123-130 (2002).
 DR EMBL; AP005371; BAC08195.1; -;
 KW Complete proteome.
 SO SEQUENCE 384 AA; 42660 MW; 68701A5F3DF424F9 CRC64;

Query Match 23.5%; Score 61; DB 16; Length 384;
 Best Local Similarity 34.1%; Pred. No. 8.6;
 Matches 15; Conservative 8; Mismatches 19; Indels 2; Gaps 1;

OY 5 PRGQDSPLLORPHOIMDOGQMRHSFSAGPELLRQDKRPRSGSTG 48
 Db 130 PRMQDA--IKLPDHYLFDQAQSRABRGPIPELYRQGEVTEVTSVG 171

RESULT 7
 Q9X6T3 PRELIMINARY; PRT; 293 AA.
 AC Q9X6T3;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Glycosyltransferase.
 GN SPCG.
 OS *Streptomyces spectabilis*.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycinae; Streptomycetaceae; Streptomycetes.
 OX NCBI_TaxID=68270;
 RN (1)
 RP SEQUENCE FROM N.A.

RC STRAIN=ATCC27741;
 RA Hyun C.G., Suh J.W.;
 RT "Isolation of the gene for glycosyltransferase and N-
 RT methyltransferase from *Streptomyces spectabilis* ATCC27741, a
 RT spectinomycin producer";
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF145038; AAD28487.1; -;
 DR InterPro; IPR001173; Glyco.trans.2.
 DR Pfam; PF00535; Glycos_transf_2; 1.
 KW Transferase
 SO SEQUENCE 293 AA; 32831 MW; A52A4A0F4B5A5A6B CRC64;

Query Match 23.1%; Score 60; DB 2; Length 293;
 Best Local Similarity 41.7%; Pred. No. 8.6;
 Matches 15; Conservative 7; Mismatches 12; Indels 2; Gaps 1;

OY 1 EPGLPRQDSPLLORPHOIMDOGQMRHSFSAGPELL 36
 Db 236 EPWVCTGRDAPL--PGRLLQQRPRRHARGGRDRL 269

RESULT 8
 Q9RT11 PRELIMINARY; PRT; 333 AA.
 AC Q9RT11;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE ADP-ribosylglycohydrolase, putative.
 GN DR1958.
 OS *Deinococcus radiodurans*.
 OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
 OC Deinococcaceae; Deinococcus.
 OX NCBI_TaxID=1299;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=RI;
 RX MEDLINE=20036896; PubMed=10567266;
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
 RA Whittie R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
 RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
 RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
 RA Makarov K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
 RA Fraser C.M.;
 RT "Genome sequence of the radioreistant bacterium *Deinococcus*
 RT *radiodurans* RI";
 RL Science 286:1571-1577 (1999).
 DR EMBL; AB002034; AAF11509.1; -;
 DR TIGR; DR1958; -;
 DR InterPro; IPR005502; ADP_ribg1ychdro.
 DR InterPro; IPR002029; Peptidase_S8.
 DR Pfam; PF03747; ADP-ribosyl_GH; 1.
 DR PROSITE; PS00136; SUBRILASE ASP; 1.
 KW Hydrolase; Complete proteome.
 SO SEQUENCE 333 AA; 34881 MW; 47042A084B54D638 CRC64;

Query Match 23.1%; Score 60; DB 16; Length 333;
 Best Local Similarity 34.9%; Pred. No. 9.9;
 Matches 15; Conservative 7; Mismatches 11; Indels 10; Gaps 1;

OY 4 GPRGQDSPLLORPHOIMDOGQMRHSFSAG-----PELL 36
 Db 290 GARGLTSPBELVEPRLPRLRLHSHWQQRDMACTHFPPELL 332

RESULT 9
 Q9YBM1 PRELIMINARY; PRT; 345 AA.
 AC Q9YBM1;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

```

DE Hypoetical protein APE1577.
GN APE1577.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococcaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
  SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=9310339; PubMed=10382966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Anai A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki U., Kushida N., Oguchi A., Aoki K.-I., Kudota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.,
RT "Complete genome sequence of an aerobic hyper-thermophilic
  crenarchaeon, Aeropyrum pernix K1."
RL DNA Res. 6:83-101(1999).
DR EMBL; AP000062; BAA8057.1; -
KM Hypoetical protein; Complete proteome.
SQ SEQUENCE 345 AA; 38129 MW; CFID9SG2BI7E92E8 CRC64;

Query Match 22.9%; Score 59.5; DB 17; Length 345;
Best Local Similarity 31.9%; Pred. No. 12;
Matches 23; Conservative 8; Mismatches 16; Indels 25; Gaps 5;

QY 1 EPLG-----PRGQSPPLQRPQHLMQDQ-----GQM-----RHFSAGPEL-----LR 37
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 244 EPLGRLRYELGDPAPGSEPHPEVLNPDLATQSPALGRLCLHPHRAVSPPSLSQHMLR 303
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 38 --QDKPRRSGST 47
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 304 WPSPTKRKGSS 315
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

RESULT 10
O9RIX2 PRELIMINARY; PRT; 271 AA.
AC O9RIX2;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Hex (Pth) protein.
GN HHEX OR HEX (PRH).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
  SEQUENCE FROM N.A.
RC STRAIN=129; TISSUE=Liver;
RA Myint Z., Inazu T., Tanaka T., Yamada K., Keng V.W., Inoue Y.,
RA Kuriyama M., Noguchi T.,
RT "Genomic organization and promoter analysis of a mouse homeobox gene,
  Hex."
RL Submitted (Aug-1998) to the EMBL/Genbank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; AB017132; BAA76714.1; -
DR EMBL; AB017130; BAA76714.1; JOINED.
DR EMBL; AB017131; BAA76714.1; JOINED.
DR HSSP; P22808; INK3.
DR MGD; MGI:96086; Hhex.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS0071; HOMEBOX_2; 1.
DR DNA-binding; Homeobox; Nuclear protein.
SQ SEQUENCE 271 AA; 29952 MW; 4C0AD438CCAA6F4 CRC64;

```

```

Query Match 22.7%; Score 59; DB 11; Length 271;
Best Local Similarity 45.7%; Pred. No. 11;
Matches 16; Conservative 2; Mismatches 17; Indels 0; Gaps 0;

QY 1 EPLGPRGDSPLQRPQHLMQDQGRHSFSAPEL 35
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 119 DPLGKPLMSPLFLQRPFLHKKRGQVRFNSDQTVEL 153
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

RESULT 11
O9CRV1 PRELIMINARY; PRT; 280 AA.
AC O9CRV1;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hematopoietically expressed homeobox (Fragment).
GN HHEX.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
  SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai U., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Akawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gotohori T., Bono H., Kanukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochia H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Guertincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli U., Momberto N.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyszeh-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsaki S.,
RA Hayaishizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; AK014111; BAB29163.1; -
DR HSSP; P22808; INK3.
DR MGD; MGI:96086; Hhex.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS0071; HOMEBOX_2; 1.
DR DNA-binding; Homeobox; Nuclear protein.
SQ SEQUENCE 280 AA; 30771 MW; 9405DC67E1842FF0 CRC64;

Query Match 22.7%; Score 59; DB 11; Length 280;
Best Local Similarity 45.7%; Pred. No. 11;
Matches 16; Conservative 2; Mismatches 17; Indels 0; Gaps 0;

QY 1 EPLGPRGDSPLQRPQHLMQDQGRHSFSAPEL 35
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 128 DPLGKPLMSPLFLQRPFLHKKRGQVRFNSDQTVEL 162
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

RESULT 12
O9ACT4 PRELIMINARY; PRT; 381 AA.
ID O9ACT4

```

AC Q9ACT4; 09ACT4; 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 17, last sequence update)
DE Hypothetical protein SCPL217.1c.
GN SCPL217.1c OR SCPL217c.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteridae; Actinomycetales;
OC Streptomyces; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kleser T., Lake L., Murphy L., Oliver K., O'Neill S.,
RA Rabinovich E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrall B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL590464; CAC36740.1; -
KW Hypothetical protein; Plasmid; Complete proteome.
SQ SEQUENCE 381 AA; 40037 MW; 453BDB6962B7525 CRC64;

Query Match 22.7%; Score 59; DB 16; Length 381;
Best Local Similarity 36.7%; Pred. No. 16;
Matches 18; Conservative 6; Mismatches 17; Indels 8; Gaps 3;
Oy 8 QDPLQRPQHL--MDQGMHRSF---SAGPELLRQ--DKRPRSGSGS 48
Db 331 RDDPLSPRGHLRIDAGVFRHTLRLPAAGDPMLEVLNPTVARTG 379

RESULT 13
Q9NR6 PRELIMINARY; PRT; 644 AA.
AC Q9NR6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, last annotation update)
DE Phosphatidylinositol polyphosphate 5-phosphatase type IV.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20347929; PubMed=10764818;
RA Kiseleva M.V., Wilson M.P., Majerus P.W.;
RT "The isolation and characterization of a cDNA encoding phospholipid-
RT specific inositol polyphosphate 5-phosphatase.";
RL J. Biol. Chem. 275:20110-20116(2000).
DR EMBL; AF187891; AAF81404.1; -
DR InterPro; IPR005135; Exo_endo_phos.
DR Pfam; PF00372; Exo_endo_phos; 1.
DR SMART; SM00128; IPFC; 1.
SQ SEQUENCE 644 AA; 70150 MW; 1BB794522C94A49F CRC64;

Query Match 22.7%; Score 59; DB 4; Length 644;
Best Local Similarity 36.0%; Pred. No. 28;
Matches 18; Conservative 3; Mismatches 27; Indels 2; Gaps 1;
Oy 2 PLGPRGQDPLQRPQHLMDQGMHRSFSGPELL--RDKRPRSGSGS 49
Db 117 RDDPLSPRGHLRIDAGVFRHTLRLPAAGDPMLEVLNPTVARTG 379

Db 69 PIAPRPPARLRLALSLDDKGWRRRRRFRSGQEDLEARNGTSPRSGSVQS 118
RESULT 14
ID Q85275 PRELIMINARY; PRT; 276 AA.
AC Q85275;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, last annotation update)
DE P1 gene (fragment).
GN P1.
OS Potato virus Y.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Potyvirus.
OX NCBI_TaxID=12216;
RN [1]
RP SEQUENCE FROM N.A.
RA Penu T., Maki-Valkama T., Valkonen J.P.T., Koiyu K., Lehto K.,
RA Penu E.;
RT "Potato plants transformed with a potato virus Y P1 gene sequence are
RT resistant to PVY-O.";
RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; X82848; CAA58051.1; -
DR InterPro; IPR002540; Poty_P1.
DR Pfam; PF01577; Poty_P1; 1.
FT NON TER 276
SQ SEQUENCE 276 AA; 31269 MW; 00A74C8425DF2BD CRC64;

Query Match 22.5%; Score 58.5; DB 12; Length 276;
Best Local Similarity 34.5%; Pred. No. 13;
Matches 19; Conservative 8; Mismatches 15; Indels 13; Gaps 3;
Oy 1 EPLGPRG--QDPLQ---RQHLMDQGMHRSFSGPELLRQDKRPRSGSGS 49
Db 117 EPOAPRGIIHTTPRVKVKTRPIIKTEGQMDH-----LIKQVKQIMSGKRS 164

RESULT 15
Q9C2E4 PRELIMINARY; PRT; 580 AA.
AC Q9C2E4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, last annotation update)
DE Related to vacuolar protein sorting-associated protein VP55.
GN 966.210.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariiales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Aign V., Hehseisel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RA Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL513463; CAC28769.2; -
DR InterPro; IPR001683; PX.
DR Pfam; PF00787; PX; 1.
DR PROSITE; PS50195; PX; 1.
SQ SEQUENCE 580 AA; 64100 MW; 311F73EF889C7E2F CRC64;

Query Match 22.5%; Score 58.5; DB 3; Length 580;
Best Local Similarity 32.7%; Pred. No. 29;
Matches 17; Conservative 7; Mismatches 21; Indels 7; Gaps 2;
Oy 1 EPLGPRG-----QDPLQRPQHLMDQGMHRSFSGPELLRQDKRPRSG 45
Db 74 DPLGPRGASTPTATDTPLAPQPLKEQLPIKTLISGPNQIGAGARRPGGG 125

Fri Jul 25 15:45:28 2003

us-09-991-681-31.rpt

Page 6

Search completed: July 24, 2003, 11:57:09
Job time : 13.9048 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 11:58:19 ; Search time 10.1903 Seconds
(without alignments)
623.047 Million cell updates/sec

Title: US-09-991-681-30

Perfect score: 40
Sequence: 1 SPKVEKDPKRKEMENAGNKIYTMADKTIKLTMEYK 40

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1107863 seqs, 158726573 residues

Word size: 0

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: A_Geneseq_19Jun03.*

```
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	40	23	AAO19168 Human prostate-spe
2	40	100.0	40	23	AAO19168 Human prostate-spe
3	40	100.0	518	19	AAW85472 PS118 protein enco
4	40	100.0	518	23	AAO19165 Human prostate-spe
5	40	100.0	518	23	AAW85472 PS118 prostate mar
6	40	100.0	1807	22	ABG09728 Novel human diagno
7	40	100.0	1839	23	ABP64835 Human protein seq
8	40	100.0	1982	22	ABG09731 Novel human diagno
9	7	17.5	139	22	ABG16511 Novel human diagno

10	7	17.5	305	22	ABG16512 Novel human diagno
11	7	17.5	321	22	ABG16510 Novel human diagno
12	7	17.5	321	22	ABG17536 Novel human diagno
13	7	17.5	365	21	AAO32620 Arabidopsis thalia
14	7	17.5	406	21	AAO32619 Arabidopsis thalia
15	7	17.5	559	23	ABB82161 Human NOV6b protei
16	7	17.5	1029	23	AAE25161 RCH1.4 protein. U
17	7	17.5	1029	23	ABH93113 Herbicidally activ
18	7	17.5	1813	23	ABG02232 Novel human diagno
19	6	15.0	13	21	AAV70167 F. heparinum hepar
20	6	15.0	19	23	AAO17916 Human prostate sec
21	6	15.0	28	15	AAK53472 Peptide fragment o
22	6	15.0	53	23	AAO17954 Human prostate sec
23	6	15.0	54	23	AAO17955 Human prostate sec
24	6	15.0	55	23	AAO17956 Human prostate sec
25	6	15.0	56	21	AAO35275 Zea mays protein f
26	6	15.0	56	21	AAO35275 Human immune/haema
27	6	15.0	56	23	AAO17957 Human prostate sec
28	6	15.0	57	23	AAO17958 Human prostate sec
29	6	15.0	58	22	AAW83330 Human immune/haema
30	6	15.0	58	23	AAO17959 Human prostate sec
31	6	15.0	59	21	AAO35274 Zea mays protein f
32	6	15.0	59	23	AAO17960 Human prostate sec
33	6	15.0	60	23	AAO17961 Human prostate sec
34	6	15.0	61	23	AAO17962 Human prostate sec
35	6	15.0	62	23	AAO17963 Human prostate sec
36	6	15.0	63	21	AAO35273 Zea mays protein f
37	6	15.0	63	23	AAO17964 Human prostate sec
38	6	15.0	64	23	AAO17965 Human prostate sec
39	6	15.0	68	21	AAO56273 Arabidopsis thalia
40	6	15.0	83	22	ABW64091 Drosophila melanog
41	6	15.0	84	23	ABP24730 Human ORF3703 prot
42	6	15.0	88	23	ABP33511 Human ORF2484 prot
43	6	15.0	94	15	AAK47117 Complete sequence
44	6	15.0	94	23	AAO17911 Human prostate sec
45	6	15.0	95	21	AAO34834 Arabidopsis thalia

ALIGNMENTS

RESULT 1	AAO19168	AAO19168 standard; Protein; 40 AA.
ID	AAO19168;	
XX	AAO19168;	
AC	AAO19168;	
XX	AAO19168;	
DT	27-NOV-2002 (first entry)	
XX	27-NOV-2002 (first entry)	
DE	Human prostate-specific PS118 protein fragment #4.	
XX	Human; prostate; prostate-specific sequence; prostate cancer; PS118;	
KW	EST; expressed sequence tag; cytosolic; gene therapy.	
XX	EST; expressed sequence tag; cytosolic; gene therapy.	
OS	Homo sapiens.	
XX	Homo sapiens.	
PN	US2002086316-A1.	
XX	US2002086316-A1.	
PD	04-JUL-2002.	
XX	04-JUL-2002.	
PF	26-NOV-2001; 2001US-0991681.	
XX	26-NOV-2001; 2001US-0991681.	
PR	23-APR-1998; 98US-0065383.	
XX	23-APR-1998; 98US-0065383.	
PR	23-APR-1997; 97US-0842385.	
XX	23-APR-1997; 97US-0842385.	
PA	(BILU/) BILLINGEL P A.	
XX	(BILU/) BILLINGEL P A.	
PA	(COHE/) COHEN M.	
XX	(COHE/) COHEN M.	
PA	(COLP/) COLPITTS T L.	
XX	(COLP/) COLPITTS T L.	
PA	(FRIE/) FRIEDMAN P N.	
XX	(FRIE/) FRIEDMAN P N.	
PA	(GORD/) GORDAN J.	
XX	(GORD/) GORDAN J.	
PA	(GRAN/) GRANADOS E N.	
XX	(GRAN/) GRANADOS E N.	
PA	(HODG/) HODGES S C.	
XX	(HODG/) HODGES S C.	
PA	(KLAS/) KLAS M R.	
XX	(KLAS/) KLAS M R.	

PA (KRAT/) KRATOCHVIL J D.
 PA (ROBE/) ROBERTS-RAPP L.
 PA (RUSSELL/) RUSSELL J C.
 PA (STROU/) STROUPE S D.
 XX Billingsel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;
 PI Granados EN, Hodges SC, Klasse MR, Kratochvil JD, Roberts-Rapp L;
 PI Russell JC, Stroupe SD;
 XX WPI; 2002-665429/71.
 DR WPI; 2002-665429/71.
 XX
 PT Novel PS118 polypeptide for detecting, diagnosing, staging, monitoring,
 PT prognosticating, preventing, treating, or determining predisposition of
 PT individual to diseases and conditions of prostate, e.g. prostate
 PT cancer -
 PS Claim 17; Page 44; 58pp; English.
 CC The present invention relates to a number of prostate-specific sequences
 CC derived from the human PS118 gene. These can be used in the detection,
 CC monitoring and treatment of prostate diseases, particularly prostate
 CC cancer. The present sequence is a PS118 protein fragment of the
 CC invention. The coding sequences of the invention were isolated from a
 CC prostate tissue expressed sequence tag (EST) library.
 XX
 SQ Sequence 40 AA;
 Query Match 100.0%; Score 40; DB 23; Length 40;
 Best Local Similarity 100.0%; Pred. No. 1.6e-35;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SPKVEKDPSPKKEWENAGNKIYTMADKITSKLTMEYK 40
 Db 1 SPKVEKDPSPKKEWENAGNKIYTMADKITSKLTMEYK 40
 RESULT 2
 ID AAM50812 standard; Protein; 40 AA.
 XX AAM50812;
 AC AAM50812;
 DT 01-MAY-2002 (first entry)
 DE PS118 prostate marker immunogenic polypeptide.
 XX PS118; prostate; marker; prostate cancer; tumour; metastasis;
 KW benign prostatic hyperplasia; prostatic intraepithelial neoplasia;
 KW prostatitis; human; diagnosis; therapy; vaccine; immunogen.
 OS Homo sapiens.
 XX
 PN US200105758-A1.
 PD 27-DEC-2001.
 XX
 PF 23-APR-1998; 98US-0065383.
 XX
 PR 23-APR-1997; 97US-0842385.
 XX
 PA (BILL/) BILLING-MEDEL P A.
 PA (COHE/) COHEN M.
 PA (COPL/) COPLITTIS T L.
 PA (FRIE/) FRIEDMAN P N.
 PA (GORD/) GORDON J.
 PA (GRAN/) GRANADOS E N.
 PA (HODG/) HODGES S C.
 PA (KLAS/) KLASSE M R.
 PA (KRAT/) KRATOCHVIL J D.
 PA (ROBE/) ROBERTS-RAPP L.
 XX
 PI Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;
 PI Granados EN, Hodges SC, Klasse MR, Kratochvil JD, Roberts-Rapp L;

PI Russell JC, Stroupe SD;
 XX WPI; 2002-187683/24.
 DR WPI; 2002-187683/24.
 XX
 PT Detecting presence of target PS118 polynucleotide in test sample,
 PT useful for detecting, diagnosing, staging, monitoring, prognosticating,
 PT preventing or treating or determining predisposition to prostate
 PT disease -
 PS Claim 17; Page 44; 57pp; English.
 CC The present sequence is that of an immunogenic polypeptide
 CC comprising amino acids 332-371 of human prostate-specific PS118
 CC polypeptide (see AAM50809). A PS118 consensus sequence (see
 CC ABA91651) is found at least 12 times more often in prostate than
 CC in non-prostate tissue. PS118 polypeptides, polynucleotides,
 CC antibodies, agonists and inhibitors are useful for detecting,
 CC diagnosing, staging, monitoring, prognosticating, preventing and
 CC treating (including by genetic immunisation), or determining the
 CC predisposition of an individual to, diseases and conditions of the
 CC prostate, such as benign prostatic hyperplasia, prostatitis,
 CC prostatic intraepithelial neoplasia, prostate cancer, tumours and
 CC metastases. The PS118 polypeptides can be produced by expression
 CC of PS118 polynucleotides in transfected host cells, and
 CC immunogenic peptides are useful for raising PS118-specific
 CC antibodies of diagnostic use. The methods and reagents of the
 CC invention may provide an early means of detecting diseases of the
 CC prostate and may also provide new markers which can differentiate
 CC between the clinically important and unimportant prostate cancers
 CC without the use of surgery.
 XX
 SQ Sequence 40 AA;
 Query Match 100.0%; Score 40; DB 23; Length 40;
 Best Local Similarity 100.0%; Pred. No. 1.6e-35;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SPKVEKDPSPKKEWENAGNKIYTMADKITSKLTMEYK 40
 Db 1 SPKVEKDPSPKKEWENAGNKIYTMADKITSKLTMEYK 40
 RESULT 3
 ID AAM85472 standard; Protein; 518 AA.
 XX AAM85472;
 AC AAM85472;
 DT 25-MAR-2003 (updated)
 DT 25-FEB-1999 (first entry)
 DE PS118 protein encoded by consensus sequence.
 XX
 KW EST clone; PS118; prostate tumour tissue; prostatic disease; cancer.
 OS Homo sapiens.
 XX
 PN WO9848054-A1.
 PD 29-OCT-1998.
 XX
 PF 23-APR-1998; 98WO-US08239.
 XX
 PR 23-APR-1997; 97US-0842385.
 XX
 PA (ABBO/) ABBOTT LAB.
 XX
 PI Billing-medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;
 PI Granados EN, Hodges SC, Klasse MR, Kratochvil JD, Roberterrapp L;
 PI Russell JC, Stroupe SD;
 XX WPI; 1998-610000/51.
 DR N-PSDB; AAV82812.

XX New P118 nucleic acid and proteins - used for diagnosis and
PT treatment of prostatic disease, especially cancer, and also for drug
PT screening
XX
PS Claim 17; Page 93-94; 117pp; English.
XX
XX The present sequence is encoded by consensus P5118 sequence derived from
CC overlapping EST (expressed sequence tag) clones AAV82803-10. The clones
CC were identified from a cDNA library made from prostate tumour tissue.
CC Recombinant P5118 protein is used to detect P5118-specific antibodies,
CC to raise antibodies for detection of P5118 antigens, to screen for
CC specific binding agents (potential therapeutics), and to isolate specific
CC antibodies from serum. Detection of P5118 protein or nucleic acid, which
CC are prostate related, and altered or elevated in prostatic disease, is
CC used for detection, diagnosis, staging, monitoring and prognosis of
CC prostatic disease, particularly cancer, and to identify subjects at
CC risk.
CC (Updated on 25-MAR-2003 to correct PI field.)
CC
XX
SQ Sequence 518 AA;
Query Match 100.0%; Score 40; DB 19; Length 518;
Best Local Similarity 100.0%; Pred. No. 1.7e-34;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 SPKVEKDPSPRKKEMWENAGNKITYTMAADKTIISKLTMEYK 40
DB 332 SPKVEKDPSPRKKEMWENAGNKITYTMAADKTIISKLTMEYK 371
RESULT 4
ID AAO19165 standard; Protein; 518 AA.
AC AAO19165;
XX
DT 27-NOV-2002 (first entry)
XX
DE Human prostate-specific P5118 protein fragment #1.
XX
XX Human; prostate; prostate-specific sequence; prostate cancer; P5118;
KW EST; expressed sequence tag; cytostatic; gene therapy.
XX
OS Homo sapiens.
XX
EN US2002086316-A1.
XX
PD 04-JUL-2002.
XX
PF 26-NOV-2001; 2001US-0991681.
XX
PR 23-APR-1998; 98US-0065383.
PR 23-APR-1997; 97US-0842385.
XX
XX (BILL/) BILLINGEL P A.
PA (COHE/) COHEN M.
PA (COLP/) COLPITTS T L.
PA (FRIE/) FRIEDMAN P N.
PA (GORD/) GORDAN J.
PA (GRAN/) GRANADOS E N.
PA (HODG/) HODGES S C.
PA (KLAS/) KLAS M R.
PA (KRAT/) KRATOCHVIL J D.
PA (ROBE/) ROBERTS-RAPP L.
PA (RUS/) RUSSELL J C.
PA (STRO/) STROUPE S D.
XX
PI Billingsel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;
PI Granados EN, Hodges SC, Kلاس MR, Kratochvil JD, Roberts-Rapp L;
PI Russell JC, Stroupe SD;
XX WPI; 2002-665429/71.
DR

XX Novel P5118 polypeptide for detecting, diagnosing, staging, monitoring,
PT prognosticating, preventing, treating, or determining predisposition of
PT individual to diseases and conditions of prostate, e.g. prostate
PT cancer -
XX
PS Claim 17; Page 42-43; 58pp; English.
XX
XX The present invention relates to a number of prostate-specific sequences
CC derived from the human P5118 gene. These can be used in the detection,
CC monitoring and treatment of prostate diseases, particularly prostate
CC cancer. The present sequence is a P5118 protein fragment of the
CC invention. The coding sequences of the invention were isolated from a
CC prostate tissue expressed sequence tag (EST) library.
XX
SQ Sequence 518 AA;
Query Match 100.0%; Score 40; DB 23; Length 518;
Best Local Similarity 100.0%; Pred. No. 1.7e-34;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 SPKVEKDPSPRKKEMWENAGNKITYTMAADKTIISKLTMEYK 40
DB 332 SPKVEKDPSPRKKEMWENAGNKITYTMAADKTIISKLTMEYK 371
RESULT 5
ID AAM50809 standard; Protein; 518 AA.
AC AAM50809;
XX
DT 01-MAY-2002 (first entry)
XX
DE P5118 prostate marker partial sequence.
XX
XX P5118; prostate; marker; prostate cancer; tumour; metastasis;
KW benign prostatic hyperplasia; prostatic intraepithelial neoplasia;
KW prostatitis; human; diagnosis; therapy; vaccine.
XX
OS Homo sapiens.
XX
EN US2001055758-A1.
XX
PD 27-DEC-2001.
XX
PF 23-APR-1998; 98US-0065383.
PR 23-APR-1997; 97US-0842385.
XX
XX (BILL/) BILLING-MEDEL P A.
PA (COHE/) COHEN M.
PA (COLP/) COLPITTS T L.
PA (FRIE/) FRIEDMAN P N.
PA (GORD/) GORDON J.
PA (GRAN/) GRANADOS E N.
PA (HODG/) HODGES S C.
PA (KLAS/) KLAS M R.
PA (KRAT/) KRATOCHVIL J D.
PA (ROBE/) ROBERTS-RAPP L.
XX
PI Billings-Medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;
PI Granados EN, Hodges SC, Kلاس MR, Kratochvil JD, Roberts-Rapp L;
PI Russell JC, Stroupe SD;
XX WPI; 2002-187683/24.
DR N-PSDB; ABA91651.
XX
XX Detecting presence of target P5118 polynucleotide in test sample,
PT useful for detecting, diagnosing, staging, monitoring, prognosticating,
PT preventing or treating or determining predisposition to prostate
PT disease -
XX

PS Claim 17; Page 42-43; 57pp; English

The present sequence is that of a human prostate-specific PS18 polypeptide, as predicted from a partial consensus cDNA sequence (see ABA91651), and lacking from the N-terminal region. The PS18 consensus sequence is found at least 12 times more often in prostate than in non-prostate tissue. PS18 polypeptides, including derivatives of the present sequence, polynucleotides, antibodies, agonists and inhibitors are useful for detecting, diagnosing, staging, monitoring, prognosticating, preventing and treating (including by genetic immunisation), or determining the predisposition of an individual to, diseases and conditions of the prostate, such as benign prostatic hyperplasia, prostatic, prostatic intraepithelial neoplasia, prostate cancer, tumours and metastases. The PS18 polypeptides can be produced by expression of PS18 polynucleotides in transfected host cells. The methods and reagents of the invention may provide an early means of detecting diseases of the prostate and may also provide new markers which can differentiate between the clinically important and unimportant prostate cancers without the use of surgery.

SQ Sequence 518 AA;

Query Match	100.0%;	Score 40;	DB 23;	length 518;
Best Local Similarity	100.0%;	Pred. No. 1.7e-34;		
Matches 40; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

```
QY      1 SPKVEKKDPSRKKEWENAGNKIYTMAADKTISKLMTEYK 40
          |||||
Db     332 SPKVEKKDPSRKKEWENAGNKIYTMAADKTISKLMTEYK 371
```

RESULT 6
ABG09728

AC ABG09728;

DT 13-FEB-2002 (first entry)

Novel human diagnostic protein #9719.

KM Human, chromosome mapping; gene mapping; gene therapy; forensic food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2

PD 11-OCT-2001

PF 30-MAR-2001; 2001WO-US08631

PR	31-MAR-2000; 2000US-0540217.
DP	22-MIG-2000; 2000US-0640167

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

XX
XX

No. 7097-6071

PT New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity -

PS Claim 20; SEQ ID No 40087; 103pp; English

CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gen mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AB500010-AB530377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WFO
CC at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).
XX

SQ Sequence 1807 AA

Query Match	100.0%	Score 40;	DB 22;	Length 1807;
Best Local Similarity	100.0%	Pred. No. 5.2e-34;		
Matches 40;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

QY 1 SPKVEKKDPSRKKEWENAGNKIYTMADKTI SKLMTTEYK 40

Db 1621 SPKVEKKDPSRKKEWENAGNKIYTMADKTI SKLMTTEYK 1660

RESULT 7
ABP64835
ID ABP64835 standard; Protein; 1839 AA

AC ABP64835;

DT 25-FEB-2003 (first entry

DE Human protein SEQ ID 495.

KM Human, expressed sequence tag; EST;
KM hematopoietic disorder; central nervous system disease; viral infection
KM peripheral nervous system disease; infectious disease
KM immune deficiency; immune disorder; bacterial infection; allergy; cancer
KM fungal infection; autoimmune disorder; coagulation disorder; nocturnal;
KM anti-allergic; antiinflammatory; immunosuppressive; neuroprotective;
KM cytoskeletal; haemostatic; vituclide; antibacterial; fungicide;
KM immunostimulant; cerebroprotective.

OS Homo sapiens

PN WO200259260-A2

PD 01-AUG-2002

PF 16-NOV-2001; 2001WO-US42950

PR 17-NOV-2000; 2000US-0714936

PA (HYSE-) HYSEQ INC
XX

PI Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Boer E, Xiao AT, Yang Y, Wolman E, Dwanag BM.

XX
XX
WBT: 2002-590824/63

DR N-PSDB; ABQ99421
YY

PT New isolated polynucleotide, useful in reeseach, diagnostic or
PT therapeutic methods, e.g. preventing or treating disorders involving
PT aberrant protein expression or biological activity -

PS Claim 20; SEQ ID 495; 394pp; English

CC The present invention relates to novel human coding sequences
CC (ABG09268-ABG99608) and proteins (ABP64682-ABP65022). The sequences are
CC useful in therapeutic, diagnostic and research methods. The
CC polynucleotides may be used in the field of molecular biology as
CC hybridisation probes, primers for PCR, for chromosome and gene mapping,
CC for the recombinant production of protein, or in generation of anti-sense
CC DNA or RNA. The polynucleotides are useful in diagnostics as expressed
CC sequence tags (ESTs) for identifying expressed genes or for physical
CC mapping of the human genome. The proteins may be used as molecular weight
CC marker, or as nutritional sources or supplements. The proteins may be
CC used to maintain and expand cell population in a totipotential or
CC pluripotential state useful for re-engineering damaged or diseased
CC tissues, transplantation, manufacture of bio-pharmaceuticals or the
CC development of bio-sensors. The polynucleotides and proteins are useful
CC for preventing, treating or ameliorating disorders involving aberrant
CC protein expression or biological activity, e.g. haematopoietic disorders,
CC central/peripheral nervous system diseases, mechanical and traumatic
CC disorders, non-healing wounds, immune deficiencies and disorders,
CC infectious diseases caused by viral, bacterial or fungal infection,
CC autoimmune disorders, allergic reactions and conditions, coagulation
CC disorders, or cancer. The polynucleotide sequences of the invention were
CC assembled from ESTs isolated mainly by sequencing by hybridisation, and
CC in some cases, sequences obtained from one or more public databases.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIP0
CC at ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 1839 AA;

Query Match 100.0%; Score 40; DB 23; Length 1839;

Best Local Similarity 100.0%; Pred. No. 5.3e-34;

Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Gy 1 SPKVEKKDPSRKKEWENAGNKITYTMADKTIISKLMTEYK 40

Db 1653 SPKVEKKDPSRKKEWENAGNKITYTMADKTIISKLMTEYK 1692

RESULT 8
ABG09731 ID ABG09731 standard; Protein; 1982 AA.

XX AC ABG09731;

XX DT 13-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #9722.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS73918.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

XX XX Claim 20; SEQ ID No 40090; 103bp; English.

XX PS The invention relates to isolated polynucleotide (I) and
XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX CC and gene mapping, and in recombinant production of (II). The
XX CC polynucleotides are also used in diagnostics as expressed sequence tags
XX CC for identifying expressed genes. (I) is useful in gene therapy techniques
XX CC to restore normal activity of (II) or to treat disease states involving
XX CC (II). (II) is useful for generating antibodies against it, detecting or
XX CC quantitating a polypeptide in tissue, as molecular weight markers and as
XX CC a food supplement. (II) and its binding partners are useful in medical
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating
XX CC disorders involving aberrant protein expression or biological activity.
XX CC The polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC and to produce other types of data and products dependent upon DNA and
XX CC amino acid sequences. ABG00010-ABG30377 represent novel human
XX CC diagnostic amino acid sequences of the invention.
XX CC Note: The sequence data for this patent did not appear in the printed
XX CC specification, but was obtained in electronic format directly from WIP0
XX CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1982 AA;

Query Match 100.0%; Score 40; DB 22; Length 1982;

Best Local Similarity 100.0%; Pred. No. 5.7e-34;

Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Gy 1 SPKVEKKDPSRKKEWENAGNKITYTMADKTIISKLMTEYK 40

Db 1735 SPKVEKKDPSRKKEWENAGNKITYTMADKTIISKLMTEYK 1774

RESULT 9
ABG16511 ID ABG16511 standard; Protein; 139 AA.

XX AC ABG16511;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #16502.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS80698.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX Claim 20; SEQ ID No 46870; 103bp; English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
CC XX

SQ Sequence 139 AA;

Query Match 17.5%; Score 7; DB 22; Length 139;

Best Local Similarity 100.0%; Pred. No. 10;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 KTISKLM 36
|||||

Db 63 KTISKLM 69

RESULT 10

ABG16512
ID ABG16512 standard; Protein; 305 AA.

XX AC ABG16512;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #16503.

XX KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS80699.

XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity -

XX Claim 20; SEQ ID No 46871; 103bp; English.

XX The invention relates to isolated polynucleotide (I) and

CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
CC XX

SQ Sequence 305 AA;

Query Match 17.5%; Score 7; DB 22; Length 305;

Best Local Similarity 100.0%; Pred. No. 21;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 KTISKLM 36
|||||

Db 98 KTISKLM 104

RESULT 11

ABG16510
ID ABG16510 standard; Protein; 321 AA.

XX AC ABG16510;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #16501.

XX KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS80697.

XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity -

XX Claim 20; SEQ ID No 46869; 103bp; English.

XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 321 AA;

Query Match 17.5%; Score 7; DB 22; Length 321;

Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;QY 30 KTISKLM 36
|||||

Db 231 KTISKLM 237

RESULT 12

ID ABG17536 standard; Protein; 321 AA.

AC ABG17536;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #17527.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

XX MO200175067-A2.

XX 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR N-PSDB; AAS81723.

XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity

XX Claim 20; SEQ ID NO 47895; 103bp; English.

XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags

CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 321 AA;

Query Match 17.5%; Score 7; DB 22; Length 321;

Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;QY 30 KTISKLM 36
|||||

Db 231 KTISKLM 237

RESULT 13

ID AAG32620 standard; Protein; 365 AA.

AC AAG32620;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 39388.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

OS Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 21-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149420.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151428.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154029.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159638.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159884.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.

```
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 17.5%; Score 7; DB 21; Length 365;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 29 DKTISKL 35
Db 56 DKTISKL 62

RESULT 14
AAG32619
ID AAG32619 standard; Protein; 406 AA.
XX
AC AAG32619;
XX
DT 17-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 39387.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130549.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
```

```
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
```

```

PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147322.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.

```

```

PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 17.5%; Score 7; DB 21; Length 406;
Beet Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 DKTISKL 35
DB 97 DKTISKL 103

RESULT 15
ABBB2161
ID ABBB2161 standard; Protein; 559 AA.
XX
AC ABBB2161;
XX
DT 23-DEC-2002 (first entry)
XX
DE Human NOV6b protein sequence.
XX
KW NOVX; NOV6b; human; antidiabetic; anorectic; cardiac; hypotensive;
KW antiarteriosclerotic; anorectic; vinuclid; antibacterial; fungicide;
KW protozoacide; nootropic; neuroprotective; antiparinsonian; osteopathic;
KW anticonvulsant; antichronic; antiinflammatory; dermatological;
KW antiaesthetic; antileptic; gene therapy.
XX
OS Homo sapiens.
XX
FN WO200270660-A2.
XX
PD 12-SEP-2002.
XX
PF 04-MAR-2002; 2002WO-US06485.
XX
PR 02-MAR-2001; 2001US-273049P.
XX
PR 21-MAR-2001; 2001US-277791P.
XX
PR 29-MAR-2001; 2001US-279883P.
XX
PR 03-APR-2001; 2001US-281248P.
XX
PR 09-APR-2001; 2001US-282537P.
XX
PR 10-APR-2001; 2001US-282864P.
XX
PR 10-APR-2001; 2001US-282867P.
XX
PR 01-MAR-2002; 2002US-0087887.
XX
PA (CURA-) CURAGEN CORP.
XX
PA (CORT-) COR THERAPEUTICS.
XX
PI Kekuda R, Conley PB, Bin-Yang R, Hart M, Tomlinson JE, Topper JN;
PI Shinkets RA, Leach MD, Zethusen BD, Komuves L, Padigaru M;
XX
DR WPI; 2002-713441/77.
XX
DR N-PSDB; ABO79947.
XX
PT New isolated NOVX polypeptides and polynucleotides, useful for
PT preventing, diagnosing or treating NOVX-associated disorders e.g.
PT diabetes, obesity, atherosclerosis, cancer, Alzheimer's disease,
PT asthma, or infectious diseases -
XX
PS Claim 1; Page 47; 160pp; English.
XX
XX The invention relates to novel polypeptides NOVX (NOV1-NOV8) and
XX polynucleotides encoding them. The NOVX polypeptides can be expressed by
XX standard recombinant methodology. The polypeptides, polynucleotides and
XX antibodies are useful for treating, preventing or diagnosing diseases
XX such as metabolic disorders, diabetes, obesity, infectious diseases
XX (viral, bacterial, fungal, helminthic, and protozoal), anorexia, cancer,
XX cardiovascular diseases (hypertension, atherosclerosis), immune disorders
XX (osteoarthritis), neurodegenerative disorders, Alzheimer's disease,
XX Parkinson's disease, epilepsy, hematopoietic disorders, inflammatory skin
XX disorders, asthma, and various dyslipidemias. The nucleic acids and

```


CC polypeptides may also be used as targets for the identification of small
CC molecules that modulate or inhibit neurogenesis, cell differentiation,
CC cell proliferation, haematopoiesis, wound healing and angiogenesis, in
CC gene therapy, in generation of antibodies that bind immunospecifically to
CC NOVX substances for use in therapeutic or diagnostic methods. The nucleic
CC acids are further used as hybridization probes, in chromosome mapping,
CC tissue typing, preventive medicine, and pharmacogenomics. The present
CC sequence represents a human NOV6b polypeptide.

XX
SQ Sequence 559 AA;

Query March 17.5%; Score 7; DB 23; Length 559;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 30 KTISKLM 36
Db 11 KTISKLM 17

Search completed: July 24, 2003, 12:21:18
Job time : 11.1903 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6.
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 12:18:29 ; Search time 4.62665 Seconds
(without alignments)
365.801 Million cell updates/sec

Title: US-09-991-681-30

Percent score: 40
Sequence: 1 SPKYEKDPKSRKEMWENAGNKIYTMADKTISKLTMEYK 40

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310858 residues

Word size: 0

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: Issued Patents AA.*

1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PTCUS.COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	40	4	US-09-065-383-30
2	40	100.0	518	4	US-09-065-383-27
3	7	17.5	1081	4	US-09-394-272-4
4	6	15.0	28	1	US-07-899-535A-4
5	6	15.0	94	1	US-07-899-535A-1
6	6	15.0	157	4	US-09-107-532A-4959
7	6	15.0	212	4	US-09-328-352-6982
8	6	15.0	333	4	US-09-107-532A-4041
9	6	15.0	384	1	US-07-783-706-2
10	6	15.0	384	2	US-08-445-342A-2
11	6	15.0	384	5	US-09-066-481-2
12	6	15.0	384	5	PCT-US92-09124-2
13	6	15.0	405	3	US-09-537-357-54
14	6	15.0	408	4	US-09-328-352-4855
15	6	15.0	425	4	US-09-134-001C-5654
16	6	15.0	448	3	US-09-537-357-55
17	6	15.0	465	3	US-09-537-357-52
18	6	15.0	468	3	US-09-537-357-53
19	6	15.0	515	4	US-09-107-532A-5317
20	5	12.5	10	1	US-08-585-197A-53
21	5	12.5	11	2	US-08-445-342A-11
22	5	12.5	11	3	US-08-974-899-13
23	5	12.5	11	3	US-09-066-481-36
24	5	12.5	12	5	PCT-US93-05701-16
25	5	12.5	14	2	US-08-609-271-7
26	5	12.5	14	2	US-08-188-374-7
27	5	12.5	15	3	US-09-066-481-9

28	5	12.5	25	2	US-08-433-133-79	Sequence 79, Appl
29	5	12.5	25	2	US-08-433-133-80	Sequence 80, Appl
30	5	12.5	25	2	US-08-433-133-81	Sequence 81, Appl
31	5	12.5	25	2	US-08-433-133-82	Sequence 82, Appl
32	5	12.5	25	2	US-08-433-133-83	Sequence 83, Appl
33	5	12.5	25	2	US-08-433-133-84	Sequence 84, Appl
34	5	12.5	25	2	US-08-433-133-85	Sequence 85, Appl
35	5	12.5	25	2	US-08-433-133-86	Sequence 86, Appl
36	5	12.5	25	2	US-08-433-133-87	Sequence 87, Appl
37	5	12.5	25	2	US-08-433-133-88	Sequence 88, Appl
38	5	12.5	25	3	US-08-860-904-10	Sequence 10, Appl
39	5	12.5	25	3	US-08-860-904-15	Sequence 15, Appl
40	5	12.5	25	3	US-08-860-904-16	Sequence 16, Appl
41	5	12.5	25	3	US-08-860-904-17	Sequence 17, Appl
42	5	12.5	25	4	US-08-884-866A-17	Sequence 6, Appl
43	5	12.5	26	1	US-07-626-589-6	Sequence 11, Appl
44	5	12.5	26	1	US-08-323-444A-11	Sequence 6, Appl
45	5	12.5	26	1	US-08-236-410-6	

ALIGNMENTS

RESULT 1
US-09-065-383-30
Sequence 30, Application US/09065383
Patent No. 6391543
GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GORDON, JULIAN
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KLAAS, MICHAEL R.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: ROBERTS-RAPP, LISA
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/065,383
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/842,385
FILING DATE: 23-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6084, US, P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6391543e
US-09-065-383-30

Query Match 100.0%; Score 40; DB 4; Length 40;
Best Local Similarity 100.0%; Pred. No. 3.6e-36;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SPKVEKDPSPKKEWENAGNKIYTMADKTIKSLMTEYK 40
DB 1 SPKVEKDPSPKKEWENAGNKIYTMADKTIKSLMTEYK 40

RESULT 2
US-09-065-383-27

Sequence 27, Application US/09065383
Patent No. 6391543

GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GORDON, JULIAN
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KLAS, MICHAEL R.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: ROBERTS-RAPP, LISA
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
TITLE OF INVENTION: FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/065,383
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/842,385
FILING DATE: 23-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6084.US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 518 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6391543e
US-09-065-383-27

Query Match 100.0%; Score 40; DB 4; Length 518;
Best Local Similarity 100.0%; Pred. No. 3.7e-35;

Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SPKVEKDPSPKKEWENAGNKIYTMADKTIKSLMTEYK 40
DB 332 SPKVEKDPSPKKEWENAGNKIYTMADKTIKSLMTEYK 371

RESULT 3
US-09-394-272-4

Sequence 4, Application US/09394272
Patent No. 6472588

GENERAL INFORMATION:
APPLICANT: Haigler, Candace H.
APPLICANT: Holaday, A. Scott
TITLE OF INVENTION: TRANSGENIC FIBER PRODUCING PLANTS WITH INCREASED
TITLE OF INVENTION: EXPRESSION OF SUCROSE PHOSPHATE SYNTHASE
FILE REFERENCE: 201304/1000
CURRENT APPLICATION NUMBER: US/09/394,272
CURRENT FILING DATE: 1999-09-10
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 1081
TYPE: PRT
ORGANISM: Craterostigma plantagineum
US-09-394-272-4

Query Match 17.5%; Score 7; DB 4; Length 1081;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 KTIKSLM 36
DB 903 KTIKSLM 909

RESULT 4
US-07-899-535A-4

Sequence 4, Application US/07899535A
Patent No. 5428011

GENERAL INFORMATION:
APPLICANT: Sheeh, Anil R.
APPLICANT: Garde, Seema
APPLICANT: Panchal, Chandra J.
TITLE OF INVENTION: Pharmaceutical Preparations For
TITLE OF INVENTION: Inhibiting Tumours Associated With Prostate
TITLE OF INVENTION: Adenocarcinoma, Stomach Cancer and Breast Cancer.
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Mr. George Loud
STREET: 2001 Jefferson Davis Highway, Suite 306
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/899,535A
FILING DATE: 16-JUN-1992
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Loud, George A.
REGISTRATION NUMBER: 25,814
REFERENCE/DOCKET NUMBER: 8&B-A835
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-415-0960
TELEFAX: 703-415-0962
TELEX: 24 8614
INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
US-07-899-535A-4

Query Match 15.0%; Score 6; DB 1; Length 28;
Best Local Similarity 100.0%; Pred. No. 7.4;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 VEKDP 9
|||||
Db 12 VEKDP 17

RESULT 5
US-07-899-535A-1
Sequence 1, Application US/07899535A
Patent No. 5428011
GENERAL INFORMATION:
APPLICANT: Sheth, Anil R.
APPLICANT: Garde, Seema
APPLICANT: Panchal, Chandra J.
TITLE OF INVENTION: Pharmaceutical Preparations For
TITLE OF INVENTION: Inhibiting Tumours Associated With Prostate
TITLE OF INVENTION: Adenocarcinoma, Stomach Cancer and Breast Cancer.
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSER: Mr. George Loud
STREET: 2001 Jefferson Davis Highway, Suite 306
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/899,535A
FILING DATE: 16-JUN-1992
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Loud, George A.
REGISTRATION NUMBER: 25,814
REFERENCE/DOCKET NUMBER: S&B-A835
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-415-0960
TELEFAX: 703-415-0962
TELEX: 24 8614
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 94 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-07-899-535A-1

Query Match 15.0%; Score 6; DB 1; Length 94;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 VEKDP 9
|||||
Db 78 VEKDP 83

RESULT 6

US-09-107-532A-4959
Sequence 4959, Application US/09107532A
Patent No. 6583275

GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354

COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 4959:
SEQUENCE CHARACTERISTICS:
LENGTH: 157 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc_feature
LOCATION: (B) LOCATION 1..157
SEQUENCE DESCRIPTION: SEQ ID NO: 4959:
US-09-107-532A-4959

Query Match 15.0%; Score 6; DB 4; Length 157;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 24 YTMAD 29
|||||
Db 45 YTMAD 50

RESULT 7
US-09-328-352-6982
Sequence 6982, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 6982
LENGTH: 212

TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-6982

Query Match 15.0%; Score 6; DB 4; Length 212;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 ISKMT 37
DB 201 ISKMT 206

RESULT 8
US-09-107-532A-4041
Sequence 4041, Application US/09107532A

PATENT NO. 6583275

GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 7310

CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION

STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts

COUNTRY: USA
ZIP: 02354

COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660

COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A

FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598

FILING DATE: 14 May 1998

APPLICATION NUMBER: 60/051571

FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke

REGISTRATION NUMBER: 40,489

REFERENCE/DOCKET NUMBER: GTC-012

TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-8277

TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 4041:
SEQUENCE CHARACTERISTICS:
LENGTH: 333 amino acids

TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
HYPOTHETICAL: YES

ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium

FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...333

SEQUENCE DESCRIPTION: SEQ ID NO: 4041:
US-09-107-532A-4041

QY 2 PVEKK 7
DB 85 PVEKK 90

Query Match 15.0%; Score 6; DB 4; Length 333;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 9
US-07-783-706-2
Sequence 2, Application US/07783706

PATENT NO. 5714376

GENERAL INFORMATION:
APPLICANT: Sasisekharan, Rammach

APPLICANT: Moreman, Kelley L.

APPLICANT: Cooney, Charles L.

APPLICANT: Langer, Robert S.

TITLE OF INVENTION: The Heparinase Gene from Flavobacterium

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:
ADDRESSEE: Kilpatrick & Cody

STREET: 100 Peachtree Street, Suite 3100

CITY: Atlanta
STATE: Georgia

COUNTRY: US
ZIP: 30303

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/783,706

FILING DATE: 19911023

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patricia L.

REGISTRATION NUMBER: 31,284

REFERENCE/DOCKET NUMBER: MIT5546

TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-572-6555

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 384 amino acids

TYPE: AMINO ACID
STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA
HYPOTHETICAL: YES

ANTI-SENSE: YES
FRAGMENT TYPE: N-terminal

ORIGINAL SOURCE:
ORGANISM: Flavobacterium heparinum

US-07-783-706-2

Query Match 15.0%; Score 6; DB 1; Length 384;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 KVEKD 8
DB 205 KVEKD 210

RESULT 10
US-08-445-342A-2
Sequence 2, Application US/08445342A

PATENT NO. 5810726

GENERAL INFORMATION:
APPLICANT: Sasisekharan, Rammach

APPLICANT: Moreman, Kelley L.

APPLICANT: Cooney, Charles L.

APPLICANT: Langer, Robert S.

TITLE OF INVENTION: The Heparinase gene from Flavobacterium

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:
ADDRESSEE: Patricia L. Pabst

STREET: 2800 One Atlantic Center, 1201 West Peachtree Street
CITY: Atlanta
STATE: GA
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/445,342A
FILING DATE: 19-MAY-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/783,706
FILING DATE: 23-OCT-1991
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: MIT 5546 div
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 873-8795
TELEFAX: (404) 873-8794
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 384 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-445-342A-2

Query Match 15.0%; Score 6; DB 2; Length 384;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 KVEKDD 8
|||||
DB 205 KVEKDD 210

RESULT 11
US-09-066-481-2
Sequence 2, Application US/09066481B
Patent No. 6217863
GENERAL INFORMATION:
APPLICANT: GODAVARTI, RANGANATHAN
APPLICANT: SASISEKHARAN, RAMNATH
APPLICANT: ERNST, STEFFAN
APPLICANT: GANESH VENKATARAMAN
APPLICANT: COONEY, CHARLES L
APPLICANT: LANGER, ROBERT
TITLE OF INVENTION: RATIONALLY DESIGNED POLYSACCHARIDE
TITLE OF INVENTION: LYSASES DERIVED FROM HEPARINASE I
FILE REFERENCE: M0656/7038/HCL
CURRENT APPLICATION NUMBER: US/09/066,481B
CURRENT FILING DATE: 1999-01-19
EARLIER APPLICATION NUMBER: US 60/008,069
EARLIER FILING DATE: 1995-10-30
NUMBER OF SEQ ID NOS: 38
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 384
TYPE: PRT
ORGANISM: Flavobacterium Heparinum
US-09-066-481-2

Query Match 15.0%; Score 6; DB 3; Length 384;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 KVEKDD 8

DB 205 KVEKDD 210
|||||

RESULT 12
PCT-US92-09124-2
Sequence 2, Application PC/TUS9209124
GENERAL INFORMATION:
APPLICANT: Massachusetts Institute, of Technology
TITLE OF INVENTION: The Heparinase Gene from Flavobacterium
TITLE OF INVENTION: Heparinum
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kilpatrick & Cody
STREET: 1100 Peachtree Street, Suite 2800
CITY: Atlanta
STATE: Georgia
COUNTRY: US
ZIP: 30309-4530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/09124
FILING DATE: 19921022
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: MITS546
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-815-6508
TELEFAX: 404-815-6555
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 384 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Flavobacterium heparinum
PCT-US92-09124-2

Query Match 15.0%; Score 6; DB 5; Length 384;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 KVEKDD 8
|||||
DB 205 KVEKDD 210

RESULT 13
US-09-537-357-54
Sequence 54, Application US/09537357
Patent No. 6271018
GENERAL INFORMATION:
APPLICANT: Alan Brash
APPLICANT: Nathalie Tillet
TITLE OF INVENTION: MUSKIEON (CUCUMIS MELO) HYDROPEROXIDE
TITLE OF INVENTION: LYSASE AND USES THEREOF
FILE REFERENCE: 06027.0002
CURRENT APPLICATION NUMBER: US/09/537,357
CURRENT FILING DATE: 2000-03-29
NUMBER OF SEQ ID NOS: 56
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 54

Job time : 4.62665 secs

LENGTH: 405
TYPE: PRT
ORGANISM: Guanyule
US-09-537-357-54

Query Match
Best Local Similarity 15.0%; Score 6; DB 3; Length 405;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 KVEKDD 8
|||||
Db 85 KVEKDD 90

RESULT 14
US-09-328-352-4855
Sequence 4855, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 4855
LENGTH: 408
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-4855

Query Match
Best Local Similarity 15.0%; Score 6; DB 4; Length 408;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPKVEK 6
|||||
Db 372 SPKVEK 377

RESULT 15
US-09-134-001C-5654
Sequence 5654, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 5654
LENGTH: 425
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5654

Query Match
Best Local Similarity 15.0%; Score 6; DB 4; Length 425;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 30 KTISKU 35
|||||
Db 237 KTISKU 242

Search completed: July 24, 2003, 12:28:38

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: July 24, 2003, 12:26:04 ; Search time 6.79356 Seconds
(without alignments)
699.249 Million cell updates/sec

Title: US-09-991-681-30

Perfect score: 40
Sequence: 1 SPKVEKDPFRKKEWMENGNKIYTMADKTIKLTMEYK 40

Scoring table: OLIGO
Gapop 60.0, Gapext 60.0

Searched: 451899 seqs, 118759770 residues

Word size: 0

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*

2: /cgn2_6/ptodata/2/pubpaa/PTCT_NEW_PUB.pep:*

3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*

4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*

5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*

6: /cgn2_6/ptodata/2/pubpaa/PTCT_NEW_PUB.pep:*

7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*

8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*

9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*

10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*

11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*

12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*

13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*

14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*

15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*

16: /cgn2_6/ptodata/2/pubpaa/US10C_NEW_PUB.pep:*

17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*

18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7	17.5	1081	US-10-217-700-4	Sequence 4, Appl1
2	6	15.0	19	US-10-291-360-6	Sequence 6, Appl1
3	6	15.0	94	US-10-291-360-1	Sequence 1, Appl1
4	6	15.0	102	US-10-291-360-2	Sequence 2, Appl1
5	6	15.0	114	US-10-012-896-1003	Sequence 1003, Ap
6	6	15.0	114	US-10-205-823-271	Sequence 271, Ap
7	6	15.0	132	US-09-925-300-1027	Sequence 1027, Ap
8	6	15.0	384	US-09-802-285-2	Sequence 2, Appl1
9	6	15.0	384	US-10-291-317-2	Sequence 2, Appl1
10	6	15.0	405	US-09-884-260A-54	Sequence 54, Appl1
11	6	15.0	448	US-09-884-260A-55	Sequence 55, Appl1
12	6	15.0	465	US-09-884-260A-52	Sequence 52, Appl1
13	6	15.0	468	US-09-884-260A-53	Sequence 53, Appl1
14	6	15.0	513	US-10-156-761-13297	Sequence 13297, A
15	6	15.0	568	US-09-950-788-2	Sequence 2, Appl1

16	6	15.0	568	US-09-950-788-4	Sequence 4, Appl1
17	6	15.0	568	US-09-950-788-7	Sequence 7, Appl1
18	6	15.0	568	US-09-738-626-5622	Sequence 5622, Ap
19	6	15.0	573	US-09-815-242-4944	Sequence 4944, Ap
20	6	15.0	575	US-09-815-242-10785	Sequence 10785, A
21	6	15.0	577	US-09-815-242-13281	Sequence 13281, A
22	6	15.0	1143	US-09-924-154-14	Sequence 14, Appl1
23	6	15.0	1798	US-10-128-714-8388	Sequence 8388, Ap
24	6	15.0	1814	US-10-128-714-3388	Sequence 3388, Ap
25	6	12.5	16	US-10-020-441-3	Sequence 3, Appl1
26	5	12.5	17	US-09-884-017-2	Sequence 2, Appl1
27	5	12.5	17	US-10-225-567A-725	Sequence 725, Ap
28	5	12.5	19	US-09-864-761-39115	Sequence 39115, A
29	5	12.5	29	US-10-057-789-263	Sequence 263, Ap
30	5	12.5	29	US-10-212-628-263	Sequence 263, Ap
31	5	12.5	34	US-09-864-761-44900	Sequence 44900, A
32	5	12.5	34	US-09-864-761-41293	Sequence 41293, A
33	5	12.5	43	US-09-774-639-315	Sequence 315, Ap
34	5	12.5	43	US-09-864-761-36018	Sequence 36018, A
35	5	12.5	44	US-09-864-761-16018	Sequence 16018, A
36	5	12.5	44	US-09-864-761-46155	Sequence 46155, A
37	5	12.5	44	US-09-776-724A-289	Sequence 289, Ap
38	5	12.5	49	US-10-082-747A-14	Sequence 14, Appl1
39	5	12.5	49	US-10-082-747A-16	Sequence 16, Appl1
40	5	12.5	54	US-09-764-872-463	Sequence 463, Appl1
41	5	12.5	55	US-09-764-868-1221	Sequence 1221, Ap
42	5	12.5	64	US-09-864-761-41565	Sequence 41565, A
43	5	12.5	67	US-09-867-550-444	Sequence 444, Ap
44	5	12.5	67	US-09-764-891-4754	Sequence 4754, Ap
45	5	12.5	70	US-09-864-761-45808	Sequence 45808, A

ALIGNMENTS

RESULT 1

US-10-217-700-4

; Sequence 4, Application US/10217700

; Publication No. US20030070191A1

GENERAL INFORMATION:

APPLICANT: Haigler, Candace H.

APPLICANT: Holiday, A. Scott

TITLE OF INVENTION: TRANSGENIC FIBER PRODUCING PLANTS WITH INCREASED

TITLE OF INVENTION: EXPRESSION OF SUCROSE PHOSPHATE SYNTHASE

FILE REFERENCE: 201304/1000

CURRENT APPLICATION NUMBER: US/10/217,700

CURRENT FILING DATE: 2002-08-12

EARLIER APPLICATION NUMBER: 09/394,272

EARLIER FILING DATE: 1999-09-10

NUMBER OF SEQ ID NOS: 14

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 4

LENGTH: 1081

TYPE: PRT

ORGANISM: Craterostigma plantagineum

US-10-217-700-4

Query Match 17.5%; Score 7; DB 15; Length 1081;

Best Local Similarity 100.0%; Pred. No. 37;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 30 KTISKLM 36

Db 903 KTISKLM 909

RESULT 2

US-10-291-360-6

; Sequence 6, Application US/10291360

; Publication No. US20030119744A1

GENERAL INFORMATION:

APPLICANT: PROCYON BIOPHARMA INC.

TITLE OF INVENTION: PSP-94: Use for Treatment of Hypercalcemia and Bone metastasis

FILE REFERENCE: 06508-051-US-02
CURRENT APPLICATION NUMBER: US/10/291,360
CURRENT FILING DATE: 2002-11-08
PRIOR APPLICATION NUMBER: CA 2,361,736
PRIOR FILING DATE: 2001-11-08
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 6
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Polypeptide 76-94
US-10-291-360-6

Query Match 15.0%; Score 6; DB 15; Length 19;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 VEKDP 9
DB 3 VEKDP 8

RESULT 3
US-10-291-360-1
Sequence 1, Application US/10291360
Publication No. US20030119744A1
GENERAL INFORMATION:

APPLICANT: PROCYON BIOPHARMA INC.
TITLE OF INVENTION: PSP-94: Use for Treatment of Hypercalcemia and Bone metastasis
FILE REFERENCE: 06508-051-US-02
CURRENT APPLICATION NUMBER: US/10/291,360
CURRENT FILING DATE: 2002-11-08
PRIOR APPLICATION NUMBER: CA 2,361,736
PRIOR FILING DATE: 2001-11-08
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 1
LENGTH: 94
TYPE: PRT
ORGANISM: Homo sapiens
PUBLICATION INFORMATION:
AUTHORS: Ulysbeck, M., Lindstrom, C., Welber, H., Abrahamson, P.A., Lilja, H., and
AUTHORS: Lundwall, A.
TITLE: Molecular cloning of a small prostate protein, known as beta-
TITLE: microseminoprotein, PSP94 or beta-inhibin, and demonstration of transcripts
JOURNAL: Biochem. Biophys. Res Commun.
VOLUME: 164
ISSUE: 3
PAGES: 1310-1315
DATE: 1989
DATABASE ACCESSION NUMBER: GI 111436
DATABASE ENTRY DATE: 1988-08-01
US-10-291-360-1

Query Match 15.0%; Score 6; DB 15; Length 94;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 VEKDP 9
DB 78 VEKDP 83

RESULT 4
US-10-291-360-2
Sequence 2, Application US/10291360
Publication No. US20030119744A1
GENERAL INFORMATION:
APPLICANT: PROCYON BIOPHARMA INC.
TITLE OF INVENTION: PSP-94: Use for Treatment of Hypercalcemia and Bone metastasis

FILE REFERENCE: 06508-051-US-02
CURRENT APPLICATION NUMBER: US/10/291,360
CURRENT FILING DATE: 2002-11-08
PRIOR APPLICATION NUMBER: CA 2,361,736
PRIOR FILING DATE: 2001-11-08
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 2
LENGTH: 102
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: recombinant human PSP94 (rHSPSP94) produced from yeast
US-10-291-360-2

Query Match 15.0%; Score 6; DB 15; Length 102;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 VEKDP 9
DB 86 VEKDP 91

RESULT 5
US-10-012-896-1003
Sequence 1003, Application US/10012896
Publication No. US20020183251A1
GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Devin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jjiang, Yugu
APPLICANT: Kalos, Michael D.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darick
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepster, William T.
APPLICANT: Henderson, Robert A.
APPLICANT: Hural, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
APPLICANT: Vinals de Bassols, Carlota
APPLICANT: Foy, Teresa
APPLICANT: Fanger, Gary R.
APPLICANT: Wantanabe, Yoshihiro
APPLICANT: Meagher, Madeleine Joy
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
FILE REFERENCE: 210121.427C27
CURRENT APPLICATION NUMBER: US/10/012,896
CURRENT FILING DATE: 2001-12-10
NUMBER OF SEQ ID NOS: 1011
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 1003
LENGTH: 114
TYPE: PRT
ORGANISM: Homo sapiens
US-10-012-896-1003

Query Match 15.0%; Score 6; DB 14; Length 114;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 VEKDP 9
DB 98 VEKDP 103

RESULT 6

US-10-205-823-271
 ; Sequence 271, Application US/10205823
 ; Publication No. US20030108963A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Schlegel, Robert
 ; APPLICANT: Monahan, John E.
 ; APPLICANT: Endege, Wilson O.
 ; APPLICANT: Gannavarapu, Manjula
 ; APPLICANT: Gotbacheva, Bella
 ; APPLICANT: Hoersch, Sebastian
 ; APPLICANT: Kamackar, Shubhangi
 ; APPLICANT: Monsey, Angela M.
 ; APPLICANT: Glatz, Karen
 ; APPLICANT: Zhao, Xumei
 ; APPLICANT: Anderson, Duertin
 ; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
 ; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
 ; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
 ; FILE REFERENCE: MRI-044
 ; CURRENT APPLICATION NUMBER: US/10/205,823
 ; CURRENT FILING DATE: 2002-07-25
 ; PRIOR APPLICATION NUMBER: 60/307,982
 ; PRIOR FILING DATE: 2001-07-25
 ; PRIOR APPLICATION NUMBER: 60/314,356
 ; PRIOR FILING DATE: 2001-08-22
 ; PRIOR APPLICATION NUMBER: 60/325,020
 ; PRIOR FILING DATE: 2001-09-25
 ; PRIOR APPLICATION NUMBER: 60/341,746
 ; PRIOR FILING DATE: 2001-12-12
 ; PRIOR APPLICATION NUMBER: 60/362,158
 ; PRIOR FILING DATE: 2002-03-05
 ; NUMBER OF SEQ ID NOS: 455
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 271
 ; LENGTH: 114
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-205-823-271

Query Match 15.0%; Score 6; DB 15; Length 114;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 4 VEKDP 9
 |||||
 Db 98 VEKDP 103

RESULT 7

US-09-925-300-1027
 ; Sequence 1027, Application US/09925300
 ; Patent No. US20020151681A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Craig Rosen,
 ; APPLICANT: Steve Ruben
 ; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
 ; FILE REFERENCE: PA101
 ; CURRENT APPLICATION NUMBER: US/09/925,300
 ; CURRENT FILING DATE: 2001-08-10
 ; PRIOR APPLICATION NUMBER: PCT/US00/05988
 ; PRIOR FILING DATE: 2000-03-08
 ; PRIOR APPLICATION NUMBER: 60/124,270
 ; PRIOR FILING DATE: 1999-03-12
 ; NUMBER OF SEQ ID NOS: 1890
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1027
 ; LENGTH: 132
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-925-300-1027

Query Match 15.0%; Score 6; DB 10; Length 132;
 Best Local Similarity 100.0%; Pred. No. 67;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 4 VEKDP 9
 |||||
 Db 116 VEKDP 121

RESULT 8

US-09-802-285-2
 ; Sequence 2, Application US/09802285
 ; Patent No. US20020122793A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Dongfang
 ; APPLICANT: Pojasek, Kevin
 ; APPLICANT: Shriver, Zachary
 ; APPLICANT: Holley, Kristine
 ; APPLICANT: El-Shabrawi, Yousuf
 ; APPLICANT: Venkataraman, Ganesh
 ; APPLICANT: Sasisekharan, Ram
 ; TITLE OF INVENTION: Heparinase III and Uses Thereof
 ; FILE REFERENCE: M0656/7063HCL
 ; CURRENT APPLICATION NUMBER: US/09/802,285
 ; CURRENT FILING DATE: 2001-03-08
 ; NUMBER OF SEQ ID NOS: 3
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 384
 ; TYPE: PRT
 ; ORGANISM: Pedobacter heparinus
 US-09-802-285-2

Query Match 15.0%; Score 6; DB 10; Length 384;
 Best Local Similarity 100.0%; Pred. No. 1,6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 KVEKDD 8
 |||||
 Db 205 KVEKDD 210

RESULT 9

US-10-291-337-2
 ; Sequence 2, Application US/10291337
 ; Publication No. US20030099628A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Dongfang
 ; APPLICANT: Pojasek, Kevin
 ; APPLICANT: Shriver, Zachary
 ; APPLICANT: Holley, Kristine
 ; APPLICANT: El-Shabrawi, Yousuf
 ; APPLICANT: Venkataraman, Ganesh
 ; APPLICANT: Sasisekharan, Ram
 ; TITLE OF INVENTION: Heparinase III and Uses Thereof
 ; FILE REFERENCE: M0656/7063HCL
 ; CURRENT APPLICATION NUMBER: US/10/291,337
 ; CURRENT FILING DATE: 2002-11-08
 ; NUMBER OF SEQ ID NOS: 3
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 384
 ; TYPE: PRT
 ; ORGANISM: Pedobacter heparinus
 US-10-291-337-2

Query Match 15.0%; Score 6; DB 15; Length 384;
 Best Local Similarity 100.0%; Pred. No. 1,6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 KVEKDD 8
 |||||

; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 13297
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-13297

Query Match 15.0%; Score 6; DB 15; Length 513;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 23 YTMMA 28
|||
Db 482 YTMMA 487

RESULT 15
US-09-950-788-2
; Sequence 2, Application US/09950788
; Patent No. US2002013232A1
; GENERAL INFORMATION:
; APPLICANT: MOECKEL, Bettina
; APPLICANT: HANS, Stephan
; APPLICANT: SCHISCHKA, Natalie
; APPLICANT: PFEFFERLE, Walter
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE PTSI PROTEIN
; FILE REFERENCE: 213544USO
; CURRENT APPLICATION NUMBER: US/09/950,788
; CURRENT FILING DATE: 2001-09-13
; PRIOR APPLICATION NUMBER: DE10045496.8
; PRIOR FILING DATE: 2000-09-13
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-950-788-2

Query Match 15.0%; Score 6; DB 10; Length 568;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 24 YTMMA 29
|||
Db 454 YTMMA 459

Search completed: July 24, 2003, 12:48:29
Job time : 7.79356 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: July 24, 2003, 12:21:29 / Search time 60.2635 Seconds

(without alignments)
577.701 Million cell updates/sec

Title: US-09-991-681-30

Perfect score: 40
Sequence: 1 SPKVEKDDPSRKSEWENAGNKIYTMADKTIKLMTEYK 40

Scoring table: OLIGO
Gapop 60.0, Gapext 60.0

Searched: 5580241 seqs, 870357830 residues

Word size: 0
Total number of hits satisfying chosen parameters: 5580241

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: Pending Patents_AA_Main.*

1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097A_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US097B_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US099B_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
28: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
29: /cgn2_6/ptodata/1/paa/US103_COMB.pep.*
30: /cgn2_6/ptodata/1/paa/US104_COMB.pep.*
31: /cgn2_6/ptodata/1/paa/US104_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	100.0	40	12	US-08-842-385-9
2	40	100.0	40	25	US-09-991-681-30

3	40	100.0	141	28	US-10-221-279-7783	Sequence 7783, Ap
4	40	100.0	467	12	US-08-842-385-6	Sequence 6, Appl1
5	40	100.0	518	25	US-09-991-681-27	Sequence 27, Appl1
6	40	100.0	1770	1	PCT-US03-01943-44	Sequence 44, Appl1
7	40	100.0	1770	27	US-10-144-198-44	Sequence 44, Appl1
8	40	100.0	1807	1	PCT-US01-08631-40087	Sequence 40087, A
9	40	100.0	1839	30	PCT-US01-42550-495	Sequence 495, App
10	40	100.0	1839	30	US-10-416-993-495	Sequence 495, App
11	40	100.0	1872	1	PCT-US03-04508-32	Sequence 32, Appl1
12	40	100.0	1982	1	PCT-US01-08631-40090	Sequence 40090, A
13	40	100.0	2221	1	PCT-US03-01943-30	Sequence 30, Appl1
14	40	100.0	2221	27	US-10-144-198-30	Sequence 30, Appl1
15	40	100.0	584	27	US-10-155-881-7370	Sequence 7370, Ap
16	40	100.0	102	30	US-10-424-599-248471	Sequence 248471, A
17	40	100.0	131	28	US-10-221-279-7793	Sequence 7793, Ap
18	40	100.0	139	1	PCT-US01-08631-46870	Sequence 46870, A
19	40	100.0	177	1	PCT-US02-25485-2818	Sequence 2818, Ap
20	40	100.0	177	28	US-10-219-382-2689	Sequence 2689, Ap
21	40	100.0	177	28	US-10-219-382-2689	Sequence 2689, Ap
22	40	100.0	177	31	US-60-311-261-2689	Sequence 2689, Ap
23	40	100.0	180	1	PCT-US02-25485-1696	Sequence 1696, Ap
24	40	100.0	180	28	US-10-219-382-1698	Sequence 1698, Ap
25	40	100.0	180	28	US-10-219-382-1698	Sequence 1698, Ap
26	40	100.0	180	31	US-60-311-261-1698	Sequence 1698, Ap
27	40	100.0	256	30	US-10-424-599-198871	Sequence 198871, A
28	40	100.0	305	1	PCT-US01-08631-46871	Sequence 46871, A
29	40	100.0	321	1	PCT-US01-08631-47869	Sequence 47869, A
30	40	100.0	321	1	PCT-US01-08631-47869	Sequence 47869, A
31	40	100.0	365	19	US-09-513-996A-39388	Sequence 39388, A
32	40	100.0	365	21	US-09-708-427-36315	Sequence 36315, A
33	40	100.0	406	19	US-09-513-996A-39387	Sequence 39387, A
34	40	100.0	406	21	US-09-708-427-36314	Sequence 36314, A
35	40	100.0	449	22	US-09-791-537-122625	Sequence 122625, A
36	40	100.0	559	1	PCT-US02-06485-14	Sequence 14, Appl1
37	40	100.0	559	26	US-10-087-887-14	Sequence 14, Appl1
38	40	100.0	560	19	US-09-570-581-1130	Sequence 1130, Ap
39	40	100.0	560	1	PCT-US02-31095-1	Sequence 1, Appl1
40	40	100.0	568	31	US-60-258-275-274	Sequence 274, App
41	40	100.0	828	21	US-09-708-427-24794	Sequence 24794, A
42	40	100.0	871	21	US-09-708-427-6123	Sequence 6123, Ap
43	40	100.0	952	21	US-09-708-427-6122	Sequence 6122, Ap
44	40	100.0	985	21	US-09-708-427-24793	Sequence 24793, A
45	40	100.0	1029	21	US-09-708-427-6121	Sequence 6121, Ap

ALIGNMENTS

RESULT 1
US-08-842-385-9
Sequence 9, Application US/08042385
GENERAL INFORMATION:
APPLICANT: Russell, John
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL FOR DETECTING DISEASE OF THE PROSTATE
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/842,385
FILING DATE:
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Portembski, Priscilla E
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 6084.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/937-6365
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: None
US-08-842-385-9

Query Match 100.0%; Score 40; DB 12; Length 40;
Best Local Similarity 100.0%; Pred. No. 2.2e-34;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPKVEKKDPSRRKKEWENAGNKIYTMADKTIISKLMTEYK 40
Db 1 SPKVEKKDPSRRKKEWENAGNKIYTMADKTIISKLMTEYK 40

RESULT 2
US-09-991-681-30
Sequence 30, Application US/09991681
GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA
COHEN, MAURICE
COLPITTS, TRACEY L.
FRIEDMAN, PAULA N.
GORDON, JULIAN
GRANADOS, EDWARD N.
HODGES, STEVEN C.
KLASS, MICHAEL R.
KRATOCHVIL, JON D.
ROBERTS-RAPP, LISA
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESS: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/991,681
FILING DATE: 26-Nov-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/065,383
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6084.US.PI
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: None
SEQUENCE DESCRIPTION: SEQ ID NO: 30:
US-09-991-681-30

Query Match 100.0%; Score 40; DB 25; Length 40;
Best Local Similarity 100.0%; Pred. No. 2.2e-34;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPKVEKKDPSRRKKEWENAGNKIYTMADKTIISKLMTEYK 40
Db 1 SPKVEKKDPSRRKKEWENAGNKIYTMADKTIISKLMTEYK 40

RESULT 3
US-10-221-279-7783
Sequence 7783, Application US/10221279
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
FILE REFERENCE: 21272-046
CURRENT APPLICATION NUMBER: US/10/221,279
PRIOR FILING DATE: 2002-09-06
PRIOR APPLICATION NUMBER: 09/574,454
PRIOR FILING DATE: 2000-05-19
PRIOR APPLICATION NUMBER: 09/519,705
PRIOR FILING DATE: 2000-03-07
NUMBER OF SEQ ID NOS: 12360
SOFTWARE: Custom
SEQ ID NO 7783
LENGTH: 141
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(141)
OTHER INFORMATION: Xaa = any amino acid or nothing
US-10-221-279-7783

Query Match 100.0%; Score 40; DB 28; Length 141;
Best Local Similarity 100.0%; Pred. No. 7.2e-34;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPKVEKKDPSRRKKEWENAGNKIYTMADKTIISKLMTEYK 40
Db 96 SPKVEKKDPSRRKKEWENAGNKIYTMADKTIISKLMTEYK 135

RESULT 4
US-08-842-385-6
Sequence 6, Application US/08842385
GENERAL INFORMATION:
APPLICANT: Russell, John
COLPITTS, TRACEY
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
FOR DETECTING DISEASE OF THE PROSTATE
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESS: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/842,385
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Porembek, Priscilla E
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 6084.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/937-6365
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: None
US-08-842-385-6

Query Match 100.0%; Score 40; DB 12; Length 467;
Best Local Similarity 100.0%; Pred. No. 2,3e-33;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPKVEKKDPSRKKEWENAGNKIYTMADKTIISKLTMEYK 40
Db 281 SPKVEKKDPSRKKEWENAGNKIYTMADKTIISKLTMEYK 320

RESULT 5

US-09-991-681-27
Sequence 27, Application US/09991681
GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA

COHEN, MAURICE
COLPITTS, TRACEY L.
FRIEDMAN, PAULA N.
GORDON, JULIAN
GRANADOS, EDWARD N.
HODGES, STEVEN C.
KLASS, MICHAEL R.
KRATOCHVIL, JON D.
ROBERTS-RAPP, LISA

TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
FOR DETECTING DISEASES OF THE PROSTATE

NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories

STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA

ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette

OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/991,681
FILING DATE: 26-Nov-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/065,383
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441

REFERENCE/DOCKET NUMBER: 6084.US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 518 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: None
SEQUENCE DESCRIPTION: SEQ ID NO: 27:
US-09-991-681-27

Query Match 100.0%; Score 40; DB 25; Length 518;
Best Local Similarity 100.0%; Pred. No. 8,2e-33;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPKVEKKDPSRKKEWENAGNKIYTMADKTIISKLTMEYK 40
Db 332 SPKVEKKDPSRKKEWENAGNKIYTMADKTIISKLTMEYK 371

RESULT 6

PCT-US03-01943-44
Sequence 44, Application PCT/US0301943
GENERAL INFORMATION:
APPLICANT: ORIGENE TECHNOLOGIES INC

TITLE OF INVENTION: CANCER GENES
FILE REFERENCE: 3U 9U 901 PCT
CURRENT APPLICATION NUMBER: PCT/US03/01943
CURRENT FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: US 10/054,935
PRIOR FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: US 60/356,130
PRIOR FILING DATE: 2002-02-14
PRIOR APPLICATION NUMBER: US 10/102,946
PRIOR FILING DATE: 2002-03-22

PRIOR APPLICATION NUMBER: US 10/117,229
PRIOR FILING DATE: 2002-04-08
PRIOR APPLICATION NUMBER: US 10/144,198
PRIOR FILING DATE: 2002-05-14
PRIOR APPLICATION NUMBER: US 10/197,824
PRIOR FILING DATE: 2002-07-19
NUMBER OF SEQ ID NOS: 102
SOFTWARE: PatentIn version 3.1

SEQ ID NO 44
LENGTH: 1770
TYPE: PRT
ORGANISM: Homo sapiens

PCT-US03-01943-44

Query Match 100.0%; Score 40; DB 1; Length 1770;
Best Local Similarity 100.0%; Pred. No. 8,2e-33;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPKVEKKDPSRKKEWENAGNKIYTMADKTIISKLTMEYK 40
Db 1584 SPKVEKKDPSRKKEWENAGNKIYTMADKTIISKLTMEYK 1623

RESULT 7

US-10-144-198-44
Sequence 44, Application US/10144198
GENERAL INFORMATION:
APPLICANT: Origene Technologies Inc

TITLE OF INVENTION: Regulated Cance Genes
FILE REFERENCE: 9U 105 R1
CURRENT APPLICATION NUMBER: US/10/144,198
CURRENT FILING DATE: 2002-05-14
NUMBER OF SEQ ID NOS: 44
SOFTWARE: PatentIn version 3.0

SEQ ID NO 44
LENGTH: 1770
TYPE: PRT
ORGANISM: Homo sapiens
US-10-144-138-44

Query Match 100.0%; Score 40; DB 27; Length 1770;
Best Local Similarity 100.0%; Pred. No. 8.2e-33;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SPKVEKDPSSRKKEWENAGNKIYTMADKTIISKLTMEYK 40
DB 1584 SPKVEKDPSSRKKEWENAGNKIYTMADKTIISKLTMEYK 1623

RESULT 8
PCT-US01-08631-40087
Sequence 40087, Application PC/TUS0108631
GENERAL INFORMATION:

APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 21272-049
CURRENT APPLICATION NUMBER: PCT/US01/08631
CURRENT FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 09/540,217
PRIOR FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: 09/649,167
PRIOR FILING DATE: 2000-08-23
NUMBER OF SEQ ID NOS: 60736
SOFTWARE: Custom
SEQ ID NO 40087
LENGTH: 1807
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (48)-(62)
OTHER INFORMATION: Fusion glycoprotein F0 domain identified by eMATRIX,
OTHER INFORMATION: accession number PF00523D, p-value=7.188e-10, raw score of 11.39
LOCATION: (941)-(950)
OTHER INFORMATION: Helper component proteinase domain identified by Pfam,
OTHER INFORMATION: accession name Peptidase_C6, E-value=0.0056, Pfam score of 7.9
PCT-US01-08631-40087

Query Match 100.0%; Score 40; DB 1; Length 1807;
Best Local Similarity 100.0%; Pred. No. 8.4e-33;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SPKVEKDPSSRKKEWENAGNKIYTMADKTIISKLTMEYK 40
DB 1621 SPKVEKDPSSRKKEWENAGNKIYTMADKTIISKLTMEYK 1660

RESULT 9
PCT-US01-42950-495
Sequence 495, Application PC/TUS0142950
GENERAL INFORMATION:

APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 21272-096
CURRENT APPLICATION NUMBER: PCT/US01/42950
CURRENT FILING DATE: 2001-11-16
PRIOR APPLICATION NUMBER: 09/714,936
PRIOR FILING DATE: 2000-11-17
NUMBER OF SEQ ID NOS: 682
SOFTWARE: PatentIn version 3.0
SEQ ID NO 495
LENGTH: 1839
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US01-42950-495

Query Match 100.0%; Score 40; DB 1; Length 1839;
Best Local Similarity 100.0%; Pred. No. 8.5e-33;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SPKVEKDPSSRKKEWENAGNKIYTMADKTIISKLTMEYK 40
DB 1653 SPKVEKDPSSRKKEWENAGNKIYTMADKTIISKLTMEYK 1692

RESULT 10
US-10-416-993-495
Sequence 495, Application US/10416993
GENERAL INFORMATION:

APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 21272-096
CURRENT APPLICATION NUMBER: US/10/416,993
CURRENT FILING DATE: 2003-11-16
PRIOR APPLICATION NUMBER: 09/714,936
PRIOR FILING DATE: 2000-11-17
NUMBER OF SEQ ID NOS: 682
SOFTWARE: PatentIn version 3.0
SEQ ID NO 495
LENGTH: 1839
TYPE: PRT
ORGANISM: Homo sapiens
US-10-416-993-495

Query Match 100.0%; Score 40; DB 30; Length 1839;
Best Local Similarity 100.0%; Pred. No. 8.5e-33;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SPKVEKDPSSRKKEWENAGNKIYTMADKTIISKLTMEYK 40
DB 1653 SPKVEKDPSSRKKEWENAGNKIYTMADKTIISKLTMEYK 1692

RESULT 11
PCT-US03-04508-32
Sequence 32, Application PC/TUS0304508
GENERAL INFORMATION:

APPLICANT: GATELY, DENNIS
TITLE OF INVENTION: PROSTATE SPECIFIC GENES AND THE USE THEREOF IN DESIGN
FILE REFERENCE: 037003/0301985
CURRENT APPLICATION NUMBER: PCT/US03/04508
CURRENT FILING DATE: 2003-02-19
PRIOR APPLICATION NUMBER: 60/357,140
PRIOR FILING DATE: 2002-02-19
PRIOR APPLICATION NUMBER: 60/396,082
PRIOR FILING DATE: 2002-07-17
PRIOR APPLICATION NUMBER: 60/386,759
PRIOR FILING DATE: 2002-06-10
NUMBER OF SEQ ID NOS: 93
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 32
LENGTH: 1872
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US03-04508-32

Query Match 100.0%; Score 40; DB 1; Length 1872;
Best Local Similarity 100.0%; Pred. No. 8.7e-33;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SPKVEKDPSSRKKEWENAGNKIYTMADKTIISKLTMEYK 40
DB 1686 SPKVEKDPSSRKKEWENAGNKIYTMADKTIISKLTMEYK 1725

RESULT 12
PCT-US01-08631-40090

```
; Sequence 40090, Application PC/TUS0108631
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-049
; CURRENT APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 40090
; LENGTH: 1982
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (11)..(25)
; OTHER INFORMATION: Fusion glycoprotein P0 domain identified by eMATRIX.
; OTHER INFORMATION: accession number PF00523D, p-value=7.188e-10, raw score of 11.39
; NAME/KEY: DOMAIN
; LOCATION: (1065)..(1074)
; OTHER INFORMATION: Helper component proteinase domain identified by Pfam,
; OTHER INFORMATION: accession name Pepsidase_C6, E-value=0.0056, Pfam score of 7.9
PCT-US01-08631-40090
```

```
Query Match 100.0%; Score 40; DB 1; Length 1982;
Best Local Similarity 100.0%; Pred. No. 9.2e-33;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 1 SPKVEKKDPSRKKEWENAGNKIYTMADKTIISKLMTEYK 40
Db 1735 SPKVEKKDPSRKKEWENAGNKIYTMADKTIISKLMTEYK 1774
```

```
RESULT 13
PCT-US03-01943-30
; Sequence 30, Application PC/TUS0301943
; GENERAL INFORMATION:
; APPLICANT: ORIGENE TECHNOLOGIES, INC
; TITLE OF INVENTION: CANCER GENES
; FILE REFERENCE: 3U 9U 901 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/01943
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: US 10/054,935
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: US 60/356,130
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 10/102,946
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: US 10/117,229
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: US 10/144,198
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: US 10/197,824
; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30
; LENGTH: 2221
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-01943-30
```

```
Query Match 100.0%; Score 40; DB 1; Length 2221;
Best Local Similarity 100.0%; Pred. No. 1e-32;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 1 SPKVEKKDPSRKKEWENAGNKIYTMADKTIISKLMTEYK 40
Db 2035 SPKVEKKDPSRKKEWENAGNKIYTMADKTIISKLMTEYK 2074
```

```
RESULT 14
US-10-144-198-30
; Sequence 30, Application US/10144198
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies Inc
; TITLE OF INVENTION: Regulated Prostate Cance Genes
; FILE REFERENCE: 9U 105 R1
; CURRENT APPLICATION NUMBER: US/10/144,198
; CURRENT FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 30
; LENGTH: 2221
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-198-30
```

```
Query Match 100.0%; Score 40; DB 27; Length 2221;
Best Local Similarity 100.0%; Pred. No. 1e-32;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 1 SPKVEKKDPSRKKEWENAGNKIYTMADKTIISKLMTEYK 40
Db 2035 SPKVEKKDPSRKKEWENAGNKIYTMADKTIISKLMTEYK 2074
```

```
RESULT 15
US-10-155-881-7370
; Sequence 7370, Application US/10155881
; GENERAL INFORMATION:
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Kovacic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: McIninch, Linda L.
; APPLICANT: Mufliyya, James
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(15300)J
; CURRENT APPLICATION NUMBER: US/10/155,881
; CURRENT FILING DATE: 2002-05-22
; NUMBER OF SEQ ID NOS: 37595
; SEQ ID NO 7370
; LENGTH: 584
; TYPE: PRT
; ORGANISM: Zea mays
US-10-155-881-7370
```

```
Query Match 20.0%; Score 8; DB 27; Length 584;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 7 KDPSSRKE 14
Db 494 KDPSSRKE 501
```

```
Search completed: July 24, 2003, 12:45:56
Job time : 61.2635 secs
```

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: July 24, 2003, 12:22:19 ; Search time 1.1713 Seconds
(without alignments)
140.227 Million cell updates/sec

Title: US-09-991-681-30

Perfect score: 40
Sequence: 1 SPKVEKDPSPRKKEWENAGNKITYMADKTISKLTMEYK 40

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 41799 seqs, 4106219 residues

Word size : 0

Total number of hits satisfying chosen parameters: 41799

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Pending Patents_AA New:*
1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	100.0	1872	US-10-367-978-32	Sequence 32, Appl
2	6	15.0	568	US-10-450-055-16	Sequence 16, Appl
3	5	12.5	39	US-09-439-410A-74	Sequence 74, Appl
4	5	12.5	123	US-10-273-573-9406	Sequence 9406, Ap
5	5	12.5	176	US-10-273-573-10764	Sequence 10764, A
6	5	12.5	188	US-10-374-979-99	Sequence 99, Appl
7	5	12.5	189	US-09-439-410A-47	Sequence 47, Appl
8	5	12.5	191	US-10-451-139-2	Sequence 2, Appl
9	5	12.5	208	US-10-273-573-8362	Sequence 8362, Ap
10	5	12.5	226	PCT-US03-20480-29	Sequence 29, Appl
11	5	12.5	242	US-10-450-186-49	Sequence 49, Appl
12	5	12.5	250	PCT-US02-159108-35	Sequence 35, Appl
13	5	12.5	251	PCT-US02-159108-6	Sequence 6, Appl
14	5	12.5	251	PCT-US02-159108-29	Sequence 29, Appl
15	5	12.5	251	PCT-US02-159108-31	Sequence 31, Appl
16	5	12.5	251	PCT-US02-159108-33	Sequence 33, Appl
17	5	12.5	251	PCT-US02-159108-37	Sequence 37, Appl
18	5	12.5	251	PCT-US02-159108-39	Sequence 39, Appl
19	5	12.5	277	US-10-273-573-9407	Sequence 9407, Ap
20	5	12.5	297	US-10-273-573-6718	Sequence 6718, Ap
21	5	12.5	318	US-10-273-573-6556	Sequence 6556, Ap
22	5	12.5	330	US-10-366-709-52	Sequence 52, Appl
23	5	12.5	332	PCT-US02-159108-62	Sequence 62, Appl
24	5	12.5	344	PCT-US02-159108-52	Sequence 52, Appl
25	5	12.5	348	PCT-US02-159108-54	Sequence 54, Appl
26	5	12.5	357	PCT-US02-159108-56	Sequence 56, Appl

27	5	12.5	362	US-10-098-221-7	Sequence 7, Appl
28	5	12.5	392	PCT-US02-159108-50	Sequence 50, Appl
29	5	12.5	449	PCT-US03-20751-1	Sequence 1, Appl
30	5	12.5	449	US-10-273-573-10096	Sequence 10096, A
31	5	12.5	464	US-10-029-988B-26	Sequence 26, Appl
32	5	12.5	464	US-10-032-0378-26	Sequence 26, Appl
33	5	12.5	464	US-10-273-573-10760	Sequence 10760, A
34	5	12.5	469	US-10-411-910A-80	Sequence 80, Appl
35	5	12.5	474	PCT-US02-26584-3	Sequence 3, Appl
36	5	12.5	474	US-10-461-148-1	Sequence 1, Appl
37	5	12.5	476	PCT-US02-26584-16	Sequence 16, Appl
38	5	12.5	476	US-10-461-148-9	Sequence 9, Appl
39	5	12.5	481	PCT-US03-19070-5	Sequence 5, Appl
40	5	12.5	481	US-10-463-260-5	Sequence 5, Appl
41	5	12.5	502	US-10-463-190-103	Sequence 103, App
42	5	12.5	502	US-10-463-190-104	Sequence 104, App
43	5	12.5	502	US-10-463-190-108	Sequence 108, App
44	5	12.5	502	US-10-463-190-109	Sequence 109, App
45	5	12.5	508	US-10-273-573-10246	Sequence 10246, A

ALIGNMENTS

```
RESULT 1
US-10-367-978-32
; Sequence 32, Application US/10367978
; GENERAL INFORMATION:
; APPLICANT: GATELY, DENNIS
; TITLE OF INVENTION: PROSTATE SPECIFIC GENES AND THE USE THEREOF IN DESIGN
; FILE REFERENCE: 037003-0301988
; CURRENT APPLICATION NUMBER: US/10/367,978
; PRIOR FILING DATE: 2003-02-19
; PRIOR APPLICATION NUMBER: 60/357,140
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: 60/396,082
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 60/386,759
; PRIOR FILING DATE: 2002-06-10
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 1872
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-367-978-32

Query Match          100.0%; Score 40; DB 6; Length 1872;
Best Local Similarity 100.0%; Pred. No. 2,3e-33;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SPKVEKDPSPRKKEWENAGNKITYMADKTISKLTMEYK 40
Db 1686 SPKVEKDPSPRKKEWENAGNKITYMADKTISKLTMEYK 1725

RESULT 2
US-10-450-055-16
; Sequence 16, Application US/10450055
; GENERAL INFORMATION:
; APPLICANT: BASF Aktiengesellschaft
; TITLE OF INVENTION: Novel genes of Corynebacterium
; FILE REFERENCE: 936 2000
; CURRENT APPLICATION NUMBER: US/10/450,055
; PRIOR FILING DATE: 2003-06-10
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Vers. 2.0
; SEQ ID NO 16
; LENGTH: 566
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
; US-10-450-055-16
```

Query Match 15.0%; Score 6; DB 6; Length 568;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 YTMAD 29
 |||||
 DB 454 YTMAD 459

RESULT 3
 US-09-439-410A-74
 ; Sequence 74, Application US/09439410A
 ; GENERAL INFORMATION:

; APPLICANT: Cismowski, Mary
 ; APPLICANT: Duzic, Emil
 ; TITLE OF INVENTION: AGS PROTEIN AND NUCLEOTIC ACID MOLECULES AND USES THEREOF
 ; FILE REFERENCE: 1919/60388-B
 ; CURRENT APPLICATION NUMBER: US/09/439,410A
 ; CURRENT FILING DATE: 1999-11-11
 ; NUMBER OF SEQ ID NOS: 118
 ; SOFTWARE: Patent version 3.1
 ; SEQ ID NO 74
 ; LENGTH: 39
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-439-410A-74

Query Match 12.5%; Score 5; DB 5; Length 39;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 MTEYK 40
 |||||
 DB 1 MTEYK 5

RESULT 4
 US-10-273-573-9406
 ; Sequence 9406, Application US/10273573
 ; GENERAL INFORMATION:
 ; APPLICANT: Hyseq, Inc
 ; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
 ; FILE REFERENCE: 21272-066
 ; CURRENT APPLICATION NUMBER: US/10/273,573
 ; CURRENT FILING DATE: 2002-10-18
 ; PRIOR APPLICATION NUMBER: 09/522,929
 ; PRIOR FILING DATE: 2000-04-18
 ; PRIOR APPLICATION NUMBER: 09/770,160
 ; PRIOR FILING DATE: 2001-01-26
 ; NUMBER OF SEQ ID NOS: 10994
 ; SOFTWARE: Custom
 ; SEQ ID NO 9406
 ; LENGTH: 123
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-273-573-9406

Query Match 12.5%; Score 5; DB 6; Length 123;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 WENA 19
 |||||
 DB 22 WENA 26

RESULT 5
 US-10-273-573-10764
 ; Sequence 10764, Application US/10273573
 ; GENERAL INFORMATION:
 ; APPLICANT: Hyseq, Inc
 ; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES

FILE REFERENCE: 21272-066
 ; CURRENT APPLICATION NUMBER: US/10/273,573
 ; CURRENT FILING DATE: 2002-10-18
 ; PRIOR APPLICATION NUMBER: 09/522,929
 ; PRIOR FILING DATE: 2000-04-18
 ; PRIOR APPLICATION NUMBER: 09/770,160
 ; PRIOR FILING DATE: 2001-01-26
 ; NUMBER OF SEQ ID NOS: 10994
 ; SOFTWARE: Custom
 ; SEQ ID NO 10764
 ; LENGTH: 176
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-273-573-10764

Query Match 12.5%; Score 5; DB 6; Length 176;
 Best Local Similarity 100.0%; Pred. No. 79;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 KTISK 34
 |||||
 DB 49 KTISK 53

RESULT 6
 US-10-374-979-99
 ; Sequence 99, Application US/10374979
 ; GENERAL INFORMATION:
 ; APPLICANT: John P. Carulli et al.
 ; TITLE OF INVENTION: THE HIGH BONE MASS GENE OF 11q13.3
 ; FILE REFERENCE: 032796-021
 ; CURRENT APPLICATION NUMBER: US/10/374,979
 ; CURRENT FILING DATE: 2003-03-04
 ; PRIOR APPLICATION NUMBER: US 09/544,398
 ; PRIOR FILING DATE: 2000-04-05
 ; PRIOR APPLICATION NUMBER: US 09/543,771
 ; PRIOR FILING DATE: 2000-04-05
 ; PRIOR APPLICATION NUMBER: US 09/229,319
 ; PRIOR FILING DATE: 1999-01-13
 ; PRIOR APPLICATION NUMBER: US 60/071,449
 ; PRIOR FILING DATE: 1998-01-13
 ; PRIOR APPLICATION NUMBER: US 60/105,511
 ; PRIOR FILING DATE: 1998-10-23
 ; NUMBER OF SEQ ID NOS: 109
 ; SEQ ID NO 99
 ; LENGTH: 188
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-374-979-99

Query Match 12.5%; Score 5; DB 6; Length 188;
 Best Local Similarity 100.0%; Pred. No. 84;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 VEKXD 8
 |||||
 DB 174 VEKXD 178

RESULT 7
 US-09-439-410A-47
 ; Sequence 47, Application US/09439410A
 ; GENERAL INFORMATION:
 ; APPLICANT: Cismowski, Mary
 ; APPLICANT: Duzic, Emil
 ; TITLE OF INVENTION: AGS PROTEIN AND NUCLEOTIC ACID MOLECULES AND USES THEREOF
 ; FILE REFERENCE: 1919/60388-B
 ; CURRENT APPLICATION NUMBER: US/09/439,410A
 ; CURRENT FILING DATE: 1999-11-11
 ; NUMBER OF SEQ ID NOS: 118
 ; SOFTWARE: Patent version 3.1
 ; SEQ ID NO 47
 ; LENGTH: 189

TYPE: PRT
ORGANISM: Homo sapiens
US-09-439-410A-47

Query Match
Best Local Similarity 12.5%; Score 5; DB 5; Length 189;
Matches 5; Conservative 0; Pred. No. 84; Mismatches 0; Indels 0; Gaps 0;

OY 36 MTEYK 40
|||||
Db 1 MTEYK 5

RESULT 8
US-10-451-139-2
Sequence 2, Application US/10451139
GENERAL INFORMATION:

APPLICANT: WANG, HONG
APPLICANT: ZHOU, YONGMING
APPLICANT: FOWKE, LARRY C.
APPLICANT: HER MAJESTY IN RIGHT OF CANADA AS REPRESENTED BY THE MINISTER OF AGRICULTURE AND AGRIFOOD CANADA
TITLE OF INVENTION: MODULATION OF PLANT CYCLIN-DEPENDENT KINASE INHIBITOR
FILE REFERENCE: 4810-62237
CURRENT APPLICATION NUMBER: US/10/451,139
CURRENT FILING DATE: 2003-06-18
PRIOR APPLICATION NUMBER: WO PCT/CA01/01825
PRIOR FILING DATE: 2001-12-18
PRIOR APPLICATION NUMBER: US 60/255,908
PRIOR FILING DATE: 2000-12-18
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 2
LENGTH: 191
TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-10-451-139-2

Query Match
Best Local Similarity 12.5%; Score 5; DB 6; Length 191;
Matches 5; Conservative 0; Pred. No. 85; Mismatches 0; Indels 0; Gaps 0;

OY 34 KLMTTE 38
|||||
Db 141 KLMTTE 145

RESULT 9
US-10-273-573-8362
Sequence 8362, Application US/10273573
GENERAL INFORMATION:

APPLICANT: Hygea, Inc
TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 21272-066
CURRENT APPLICATION NUMBER: US/10/273,573
CURRENT FILING DATE: 2002-10-18
PRIOR APPLICATION NUMBER: 09/522,929
PRIOR FILING DATE: 2000-04-18
PRIOR APPLICATION NUMBER: 09/770,160
PRIOR FILING DATE: 2001-01-26
NUMBER OF SEQ ID NOS: 10994
SOFTWARE: Custom
SEQ ID NO: 8362
LENGTH: 208
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (64)..(87)
OTHER INFORMATION: TRANSFORMING PROTEIN P21 RAS SIGNATURE domain identified by
OTHER INFORMATION: eMATRIX, accession number PR00449C, p-value=3.250e-20, raw score
OTHER INFORMATION: 17.27

FEATURE:
NAME/KEY: DOMAIN
LOCATION: (25)..(208)
OTHER INFORMATION: Ras family domain identified by Pfam, accession name ras, E-
US-10-273-573-8362
OTHER INFORMATION: value=1.6e-82, Pfam score of 287.6

Query Match
Best Local Similarity 12.5%; Score 5; DB 6; Length 208;
Matches 5; Conservative 0; Pred. No. 91; Mismatches 0; Indels 0; Gaps 0;

OY 36 MTEYK 40
|||||
Db 21 MTEYK 25

RESULT 10
PCT-US03-20480-29
Sequence 29, Application PC/TUS0320480
GENERAL INFORMATION:

APPLICANT: INCYTE CORPORATION
APPLICANT: ELLIOTT, Vick S.
APPLICANT: MASON, Patricia M.
APPLICANT: RICHARDSON, Thomas W.
APPLICANT: LEE, Soo Yuen
APPLICANT: RAMKUMAR, Jayalaxmi
APPLICANT: KABLE, Amy E.
APPLICANT: SMARNKAR, Anita
APPLICANT: EMERLING, Brooke M.
APPLICANT: HAPALIA, April J.A.
APPLICANT: KHARE, Reena
APPLICANT: BECHA, Shanya D.
APPLICANT: MARQUIS, Joseph P.
APPLICANT: CHEN, David
APPLICANT: JIN, Pei
APPLICANT: CHANG, Hsin-Ru
APPLICANT: BULLOCK, Sean A.
APPLICANT: TRAN, Uyen K.
APPLICANT: ISON, Craig H.
TITLE OF INVENTION: ENZYMES
FILE REFERENCE: PF-1452 PCT
CURRENT APPLICATION NUMBER: PCT/US03/20480
CURRENT FILING DATE: 2003-06-26
PRIOR APPLICATION NUMBER: US 60/393,067
PRIOR FILING DATE: 2002-06-28
PRIOR APPLICATION NUMBER: US 60/393,013
PRIOR FILING DATE: 2002-06-28
NUMBER OF SEQ ID NOS: 104
SOFTWARE: PERL Program
SEQ ID NO: 29
LENGTH: 226
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No: 7511840CD1
PCT-US03-20480-29

Query Match
Best Local Similarity 12.5%; Score 5; DB 1; Length 226;
Matches 5; Conservative 0; Pred. No. 97; Mismatches 0; Indels 0; Gaps 0;

OY 15 WMEWA 19
|||||
Db 201 WMEWA 205

RESULT 11
US-10-450-186-49
Sequence 49, Application US/10450186
GENERAL INFORMATION:
APPLICANT: INCYTE CORPORATION; GRIFFIN, Jennifer A.;
APPLICANT: YAO, Monique G.; DUGGAN, Brendan M.;

```

; APPLICANT: YUE, Henry; DING, Li;
; APPLICANT: LAL, Preeti G.; LEE, Ernestine A.;
; APPLICANT: RAMKUMAR, Jayalaxmi; THANGAVELU, Kavitha;
; APPLICANT: XU, Yuming; LEE, Sally;
; APPLICANT: TANG, Y. Tom; NGUYEN, Damien B.;
; APPLICANT: WARREN, Bridget A.; HONCHELL, Cynthia D.;
; APPLICANT: GIETZEN, Kimberly J.; BAUGHN, Mariah R.;
; APPLICANT: GANDHI, Ameeta R.; ARVIZU, Chandra S.;
; APPLICANT: CHAWLA, Narinder K.; LU, Yan;
; APPLICANT: ELIOTT, Vicki S.; LU, Dzung Anna M.;
; APPLICANT: HAFALIA, April J.A.; AZIMZAI, Yalda;
; APPLICANT: KHAN, Farrah A.; TRAN, Uyen K.
; TITLE OF INVENTION: SECRETED PROTEINS
; FILE REFERENCE: PI-0345 US
; CURRENT APPLICATION NUMBER: US/10/450,186
; CURRENT FILING DATE: 2003-06-09
; PRIOR APPLICATION NUMBER: PCT/US01/48517
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255,639
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: US 60/257,852
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: US 60/260,105
; PRIOR FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: US 60/262,932
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: US 60/263,096
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: US 60/263,090
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/265,926
; PRIOR FILING DATE: 2001-02-02
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PERL Program
; SEQ ID NO 49
; LENGTH: 242
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 7486536CD1
US-10-450-186-49

```

```

Query Match          12.5%; Score 5; DB 6; Length 242;
Best Local Similarity 100.0%; Pred. No. 1.e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 SPKVE 5
        |||||
Db      139 SPKVE 143

```

```

RESULT 12
PCT-US02-15910B-35
; Sequence 35, Application PC/TUS0215910B
; GENERAL INFORMATION:
; APPLICANT: Rixon, Mark W.
; APPLICANT: Gross, Jane A.
; TITLE OF INVENTION: TACT-Immunoglobulin Fusion Proteins
; FILE REFERENCE: 01-20PC
; CURRENT APPLICATION NUMBER: PCT/US02/15910B
; CURRENT FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: 60/293,343
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-15910B-35

```

```

Query Match          12.5%; Score 5; DB 1; Length 250;

```

```

Best Local Similarity 100.0%; Pred. No. 1.e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      30 KTISK 34
        |||||
Db      138 KTISK 142

```

```

RESULT 13
PCT-US02-15910B-6
; Sequence 6, Application PC/TUS0215910B
; GENERAL INFORMATION:
; APPLICANT: Rixon, Mark W.
; APPLICANT: Gross, Jane A.
; TITLE OF INVENTION: TACT-Immunoglobulin Fusion Proteins
; FILE REFERENCE: 01-20PC
; CURRENT APPLICATION NUMBER: PCT/US02/15910B
; CURRENT FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: 60/293,343
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-15910B-6

```

```

Query Match          12.5%; Score 5; DB 1; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      30 KTISK 34
        |||||
Db      138 KTISK 142

```

```

RESULT 14
PCT-US02-15910B-29
; Sequence 29, Application PC/TUS0215910B
; GENERAL INFORMATION:
; APPLICANT: Rixon, Mark W.
; APPLICANT: Gross, Jane A.
; TITLE OF INVENTION: TACT-Immunoglobulin Fusion Proteins
; FILE REFERENCE: 01-20PC
; CURRENT APPLICATION NUMBER: PCT/US02/15910B
; CURRENT FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: 60/293,343
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-15910B-29

```

```

Query Match          12.5%; Score 5; DB 1; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      30 KTISK 34
        |||||
Db      138 KTISK 142

```

```

RESULT 15
PCT-US02-15910B-31
; Sequence 31, Application PC/TUS0215910B
; GENERAL INFORMATION:
; APPLICANT: Rixon, Mark W.
; APPLICANT: Gross, Jane A.
; TITLE OF INVENTION: TACT-Immunoglobulin Fusion Proteins

```



```

; FILE REFERENCE: 01-20PC
; CURRENT APPLICATION NUMBER: PCT/US02/15910B
; CURRENT FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: 60/293,343
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO: 31
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-15910B-31

```

```

Query Match      12.5%; Score 5; DB 1; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      30 KTISK 34
        |||||
Db      138 KTISK 142

```

```

Search completed: July 24, 2003, 12:46:19
Job time : 1.3713 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: July 24, 2003, 12:16:59 ; Search time 4.80234 Seconds
(without alignments)
801.015 Million cell updates/sec

Title: US-09-991-681-30

Perfect score: 40
Sequence: 1 SPKYEKDDPSRKEMWENAGNKIYTMADKTISKLTMEYK 40

Scoring table: OLIGO
Gapop 60.0, Gapext 60.0

Searched: 283308 seqs, 96168682 residues

Word size: 0

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: PIR_76:*

1: pir1:*\n2: pir2:*\n3: pir3:*\n4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7	17.5	406	2 T04817	hypothetical prote
2	7	17.5	406	2 T24492	hypothetical prote
3	7	17.5	736	2 T25447	hypothetical prote
4	7	17.5	913	2 D90183	ATP-dependent heli
5	7	17.5	1029	2 T05050	protein kinase hom
6	7	17.5	1037	2 B84723	hypothetical prote
7	7	17.5	1081	2 T09837	sucrose-phosphate
8	7	17.5	73	2 A15195	hypothetical prote
9	6	15.0	78	2 S52327	hypothetical prote
10	6	15.0	91	2 S41663	acylphosphatase (E
11	6	15.0	96	2 A61427	beta-microseminopr
12	6	15.0	97	2 D89792	acylphosphatase (E
13	6	15.0	98	1 OPHO	conserved hypotet
14	6	15.0	98	1 OPRG	acylphosphatase (E
15	6	15.0	98	1 OPRB	acylphosphatase (E
16	6	15.0	98	1 OPRB	acylphosphatase (E
17	6	15.0	114	2 A45667	acylphosphatase (E
18	6	15.0	116	2 AD1451	beta-microseminopr
19	6	15.0	118	2 A61647	hypothetical prote
20	6	15.0	149	2 G90395	hypothetical prote
21	6	15.0	173	1 VOECPE	hypothetical prote
22	6	15.0	173	1 E75035	fibribial protein p
23	6	15.0	208	2 D82192	conserved hypotet
24	6	15.0	210	2 H97149	hypothetical prote
25	6	15.0	224	2 A96577	hypothetical prote
26	6	15.0	227	2 C58932	channel subunit of
27	6	15.0	230	2 H86758	hypothetical prote
28	6	15.0	237	2 T77470	hypothetical prote
29	6	15.0	246	2 E90237	conserved hypotet

30	6	15.0	252	2 B32053	hypothetical prote
31	6	15.0	255	2 T42936	hypothetical prote
32	6	15.0	274	2 T25783	hypothetical prote
33	6	15.0	281	2 D72397	conserved hypotet
34	6	15.0	299	2 S53395	hypothetical prote
35	6	15.0	311	2 S23533	hypothetical prote
36	6	15.0	313	2 E91006	penicillin-binding
37	6	15.0	313	2 F85850	penicillin-binding
38	6	15.0	313	2 E64981	penicillin-binding
39	6	15.0	317	2 S75694	hypothetical prote
40	6	15.0	324	2 F69504	proliferating-cell
41	6	15.0	328	2 H71871	hypothetical prote
42	6	15.0	331	2 A86278	hypothetical prote
43	6	15.0	331	2 T25785	hypothetical prote
44	6	15.0	332	2 T00847	probable membrane
45	6	15.0	335	2 T48161	heat shock protein

ALIGNMENTS

RESULT 1

T04817
hypothetical protein F10M23.280 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jun-1999

C:Accession: T04817

R:Bevan, M.; Lechery, A.; Chetdor, F.; Krivitzky, M.; Kreis, M.; Hohelsel, J.; Mewes, H

submitted to the Protein Sequence Database, February 1999

A:Reference number: Z15385

A:Accession: T04817

A:Molecule type: DNA

A:Residues: 1-406 <BEV>

A:Cross-references: EMBL:AL035440

A:Experimental source: cultivar Columbia; BAC clone F10M23

C:Genetics:

A:Map position: 4

A:Insertions: 41/2; 73/3; 94/2; 165/1; 186/2; 210/3; 251/1; 280/2; 322/2; 358/1

A:Note: F10M23.280

Query Match 17.5%; Score 7; DB 2; Length 406;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	29	DKTISKL 35
Db	97	DKTISKL 103

RESULT 2

T24492
hypothetical protein T05A10.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000

C:Accession: T24492

R:Sluisson, J.

submitted to the EMBL Data Library, November 1995

A:Reference number: Z19898

A:Accession: T24492

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-406 <WIL>

A:Cross-references: EMBL:Z68108; PTDN:CAA92135.1; GSPDB:GN00028; CESP:T05A10.3

A:Experimental source: clone T05A10

C:Genetics:

A:Gene: CESP:T05A10.3

A:Map position: X

A:Insertions: 50/1; 76/3; 98/3; 131/1; 167/3; 242/1; 294/1; 384/3

Query Match 17.5%; Score 7; DB 2; Length 406;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 KVEKDP 9
 Db 369 KVEKDP 375

RESULT 3

T25447
 hypothetical protein B0261.1 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T25447

R:Schect, P.; Gattung, S.
 submitted to the EMBL Data Library, April 1997

A:Description: The sequence of *C. elegans* cosmid B0261.

A:Reference number: Z20036

A:Accession: T25447

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-736 <SCH>

A:Cross-references: EMBL:U97016; PIDN:AAB52350.1; GSPDB:GN00019; CESP:B0261.1

A:Experimental source: strain Bristol NZ; clone B0261

A:Gene: CESP:B0261.1

A:Map position: 1

A:introns: 99/2; 263/2; 411/3; 541/1; 573/1; 685/1

QY 32 ISKLMTE 38
 Db 287 ISKLMTE 293

RESULT 4

D90183
 ATP-dependent helicase [imported] - *Sulfolobus solfataricus*

C:Species: *Sulfolobus solfataricus*

C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001

C:Accession: D90183

R:She, O.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Ameyez, M.J.; Chan-jong, J.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, F.

submitted to Genbank, April 2001

A:Description: *Sulfolobus solfataricus* complete genome.

A:Reference number: A99139

A:Accession: D90183

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-913 <KUR>

A:Cross-references: GB:AE006641; NID:g13131545; PIDN:AAK40723.1; GSPDB:GN00155

C:Genetics:

A:Gene: SS00394

Query Match 17.5%; Score 7; DB 2; Length 913;
 Best Local Similarity 100.0%; Pred. No. 20;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PVEKKD 8
 Db 222 PVEKKD 228

RESULT 5

T05050
 protein kinase homolog M359.30 - *Arabidopsis thaliana*

C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 20-Sep-1999

C:Accession: T05050

R:Bevan, M.; Vandebol, M.; Jallet, C.; Portetelle, D.; Honeisel, J.; Mewes, H.W.; Mayer

submitted to the Protein Sequence Database, March 1999

A:Reference number: Z15396

A:Accession: T05050
 A:Molecule type: DNA
 A:Residues: 1-1029 <BEV>

A:Cross-references: EMBL:AL022223

A:Experimental source: cultivar Columbia; BAC clone M359

C:Genetics:

A:Map position: 4

A:introns: 428/2; 862/2

A:Note: M359.30

C:Superfamily: protein kinase xaz1; leucine-rich alpha-2-glycoprotein repeat homology; p

QY 5 EKDPSP 11
 Db 925 EKDPSP 931

RESULT 6

B84723
 hypothetical protein At2g31660 [imported] - *Arabidopsis thaliana*

C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C:Accession: B84723

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanNieuwen, S.E.; Umayam, L.; Tallon, L.

neus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: B84723

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1037 <STO>

A:Cross-references: GB:AE002093; NID:g4582459; PIDN:AAD24843.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g31660

A:Map position: 2

Query Match 17.5%; Score 7; DB 2; Length 1037;
 Best Local Similarity 100.0%; Pred. No. 22;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 TISKLMT 37
 Db 721 TISKLMT 727

RESULT 7

T09837
 sucrose-phosphate synthase (EC 2.4.1.14) isoform 2 - *Cratogeomys plantaginum*

C:Species: *Cratogeomys plantaginum*

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000

C:Accession: T09837

R:Ingaram, J.; Chandler, J.W.; Gallagher, L.; Salamini, F.; Bartels, D.

Plant Physiol. 115, 113-121, 1997

A:Title: Analysis of cDNA clones encoding sucrose-phosphate synthase in relation to suga1

A:Reference number: Z16874; MUID:97451773; PMID:9306694

A:Accession: T09837

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1081 <ING>

A:Cross-references: EMBL:Y11795; NID:g2190349; PIDN:CAA72491.1; PID:g2190350

A:Experimental source: ABA-treated callus

C:Genetics:

A:Gene: sps2

C:Function:

A:Description: catalyzes the formation of sucrose-6-phosphate from UDPglucose and D-fruc

A:Pathway: sucrose biosynthesis

C:Superfamily: sucrose-phosphate synthase; sucrose/sucrose-phosphate synthase homology

C:Keywords: glycosyltransferase; hexosyltransferase; sucrose biosynthesis

F,176-674/Domains: sucrose/sucrose-phosphate synthase homology <SSS>

Query Match 17.5%; Score 7; DB 2; Length 1081;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 30 KTISKL 36
|||||
Db 903 KTISKL 909

RESULT 8

AE1595
hypothetical protein lin1302 [imported] - Listeria innocua (strain Clj11262)

C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001

C:Accession: AE1595

R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurgey, O.; Entlian, K.D.; Fsihl, H.

D.; Jones, L.M.; Karet, U.

Science 294, 849-852, 2001

A:Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; Me

ok, C.; Schluter, T.; Simoes, N.; Tietze, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.

A:Title: Comparative genomics of Listeria species.

A:Reference number: AB1077; PMID:21537279; PMID:11679669

A:Accession: AE1595

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-73 <GLA>

A:Cross-references: GB:AL592022; PIDN:CA96533.1; PID:G16413775; GSPDB:GN00178

A:Experimental source: strain Clj11262

C:Genetics:

A:Gene: lin1302

Query Match 15.0%; Score 6; DB 2; Length 73;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 KVEKXD 8
|||||
Db 41 KVEKXD 46

RESULT 9

SS2327

acylphosphatase (EC 3.6.1.7), muscle - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 08-May-1995 #sequence_revision 23-Aug-1997 #text_change 07-May-1999

C:Accession: S59138; S52327

R:Chiarugi, P.; Raugel, G.; Marzocchini, R.; Flachi, T.; Ciccarelli, C.; Berti, A.; Ram

Biochem. J. 311, 567-573, 1995

A:Title: Differential modulation of expression of the two acylphosphatase isoenzymes by

A:Reference number: S59137; PMID:96033055; PMID:7487897

A:Accession: S59138

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-78 <CH2>

A:Cross-references: EMBL:X84195

A>Note: submitted to the EMBL Data Library, January 1995

C:Genetics:

A:Gene: GDB:ACYP

A:Cross-references: GDB:304769

C:Superfamily: acylphosphatase

C:Keywords: hydrolase

Query Match 15.0%; Score 6; DB 2; Length 78;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 30 KTISKL 35
|||||
Db 64 KTISKL 69

RESULT 10

S41663

beta-microseminoprotein - pig

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 31-Oct-1997

C:Accession: S41663

R:Fernlund, P.; Granberg, L.B.; Roepstorff, P.

Arch. Biochem. Biophys. 309, 70-76, 1994

A:Title: Amino acid sequence of beta-microseminoprotein from porcine seminal plasma.

A:Reference number: S41663; PMID:94161559; PMID:8117114

A:Accession: S41663

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-91 <FER>

C:Superfamily: seminal plasma protein

Query Match 15.0%; Score 6; DB 2; Length 91;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 VEKDP 9
|||||
Db 75 VEKDP 80

RESULT 11

A61427

acylphosphatase (EC 3.6.1.7), skeletal muscle - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 07-Feb-1997

C:Accession: A61427

R:Berti, A.; Tremori, E.; Pezzagli, L.; Degl'Innocenti, D.; Camici, G.; Cappugi, G.; Man

J. Protein Chem. 10, 91-102, 1991

A:Title: Rat muscle acylphosphatase: purification, amino sequence, and immunological cha

A:Reference number: A61427; PMID:91273765; PMID:1647162

A:Accession: A61427

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-96 <BER>

C:Superfamily: acylphosphatase

C:Keywords: hydrolase; skeletal muscle

Query Match 15.0%; Score 6; DB 2; Length 96;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 30 KTISKL 35
|||||
Db 82 KTISKL 87

RESULT 12

D89792

conserved hypothetical protein SA0271 [imported] - Staphylococcus aureus (strain N315)

C:Species: Staphylococcus aureus

C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001

C:Accession: D89792

R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc

ma, A.; Mizutani-Oi, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;

C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiratazu, K.

Lancet 357, 1225-1240, 2001

A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.

A:Reference number: A89758; PMID:21311952; PMID:11418146

A:Accession: D89792

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-97 <KUR>

A:Cross-references: GB:BA000018; PID:G13700197; PIDN:BA841495.1; GSPDB:GN00149

A:Experimental source: strain N315

C:Genetics:

A:Gene: SA0271

Query Match 15.0%; Score 6; DB 2; Length 97;
 Best Local Similarity 100.0%; Pred. No. 34;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SPKVEK 6
 |||||
 DB 59 SPKVEK 64

RESULT 13

QPHO

acylphosphatase (EC 3.6.1.7) Hol. skeletal muscle - horse

N.Alternate names: acylphosphate phosphohydrolase

C.Species: Equus caballus (domestic horse)

C.Date: 31-Aug-1980 #sequence_revision 31-Aug-1980 #text_change 17-May-1996

C.Accession: A01015

R.Cappugi, G.; Manao, G.; Camici, G.; Rampont, G.

J. Biol. Chem. 255, 6668-6874, 1980

A.Title: The complete amino acid sequence of horse muscle acylphosphatase.

A.Reference number: A01015; MUID:80227852; PMID:6248536

A.Accession: A01015

A.Molecule type: protein

A.Residues: 1-98 <CAP>

C.Comment: In the absence of reducing factors, the enzyme tends to dimerize by the forma

C.Superfamily: acylphosphatase and carbamoylphosphate are possible substrates of this

C.Keywords: acetylated amino end; hydrolase

F.1/Modified site: acetylated amino end (Ser) #status experimental

F.21/Disulfide bonds: interchain (partial) #status experimental

F.21/Binding site: glutathione (Cys) (covalent) (partial) #status experimental

Query Match 15.0%; Score 6; DB 1; Length 98;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 KTISKL 35
 |||||
 DB 84 KTISKL 89

RESULT 14

OPPG

acylphosphatase (EC 3.6.1.7), skeletal muscle - pig

N.Alternate names: acylphosphate phosphohydrolase

C.Species: Sus scrofa domestica (domestic pig)

C.Date: 28-May-1986 #sequence_revision 28-May-1986 #text_change 10-May-1996

C.Accession: A01016

R.Mizuno, Y.; Yamazaki, M.; Takasawa, T.; Kizaki, T.; Shiokawa, H.

J. Biochem. 97, 1135-1142, 1985

A.Title: Amino acid sequence of acylphosphatase from porcine skeletal muscle.

A.Reference number: A01016; MUID:85289095; PMID:2993259

A.Accession: A01016

A.Molecule type: protein

A.Residues: 1-98 <MIZ>

C.Comment: 1,3-Diphosphoglycerate and carbamoylphosphate are possible substrates of this

C.Superfamily: acylphosphatase

C.Keywords: acetylated amino end; hydrolase

F.1/Modified site: acetylated amino end (Ser) #status experimental

F.21/Disulfide bonds: interchain #status predicted

F.21/Binding site: glutathione (Cys) (covalent) #status predicted

Query Match 15.0%; Score 6; DB 1; Length 98;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 KTISKL 35
 |||||
 DB 84 KTISKL 89

RESULT 15

QPHO

acylphosphatase (EC 3.6.1.7), skeletal muscle - bovine

C.Species: Bos primigenius taurus (cattle)

C.Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 10-May-1996

C.Accession: A29578

R.Camici, G.; Manao, G.; Modesti, A.; Stefani, M.; Berti, A.; Cappugi, G.; Rampont, G.

Ital. J. Biochem. 35, 1-15, 1986

A.Title: The complete amino acid sequence of bovine skeletal muscle acylphosphatase.

A.Reference number: A29578; MUID:86223180; PMID:3011706

A.Accession: A29578

A.Molecule type: protein

A.Residues: 1-98 <CAM>

C.Superfamily: acylphosphatase

C.Keywords: acetylated amino end; hydrolase

F.1/Modified site: acetylated amino end (Ser) #status experimental

F.21/Disulfide bonds: interchain #status predicted

F.21/Binding site: glutathione (Cys) (covalent) #status predicted

Query Match 15.0%; Score 6; DB 1; Length 98;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 KTISKL 35
 |||||
 DB 84 KTISKL 89

Search completed: July 24, 2003, 12:27:21
 Job time : 7.80234 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using bw model

Run on: July 24, 2003, 12:15:23 ; Search time 2.63543 Seconds

(without alignments)
713.761 Million cell updates/sec

Title: US-09-991-681-30

Perfect score: 40
Sequence: 1 SPKVEKDPSPKKEWENAGNKIYTMADKTISKLTMEYK 40

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size: 0

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7	17.5	449	1	EF1A_MANES
2	7	17.5	1081	1	SPS2_CRAPL
3	6	15.0	96	1	ACYM_RAT
4	6	15.0	97	1	ACYM_MOUSE
5	6	15.0	98	1	ACYM_BOVIN
6	6	15.0	98	1	ACYM_HORSE
7	6	15.0	98	1	ACYM_HUMAN
8	6	15.0	98	1	ACYM_PIG
9	6	15.0	98	1	ACYM_RABIT
10	6	15.0	111	1	MSMB_PIG
11	6	15.0	112	1	MSBP_SAGO
12	6	15.0	114	1	MSMB_HUMAN
13	6	15.0	133	1	FLIN_BUCAP
14	6	15.0	173	1	PAPE_ECOLI
15	6	15.0	252	1	YCR2_BACTR
16	6	15.0	313	1	PBP7_ECOLI
17	6	15.0	317	1	VN35_ROTUM
18	6	15.0	322	1	TAD3_YEAS
19	6	15.0	333	1	SECF_RHOA
20	6	15.0	340	1	G3P_SULSO
21	6	15.0	348	1	EXOB_AZOBR
22	6	15.0	357	1	TCSA_LISIN
23	6	15.0	357	1	TCSA_LISMO
24	6	15.0	384	1	HEP1_PEDHE
25	6	15.0	400	1	YV6L_CAERL
26	6	15.0	404	1	Y349_MYCB
27	6	15.0	409	1	AXIA_BRABE
28	6	15.0	412	1	YEIT_ECOS7
29	6	15.0	412	1	YEIT_ECOLI
30	6	15.0	429	1	PMR2_ANTEL
31	6	15.0	435	1	EP1A_TETRY
32	6	15.0	435	1	PMR1_ANTEL
33	6	15.0	447	1	EP12_DAVCA

34	6	15.0	447	1	EF1A_MAIZE	Q41803 zea mays (m
35	6	15.0	447	1	EF1A_ORYZA	O64937 oryza sativ
36	6	15.0	447	1	EF1A_PEA	Q41011 pisum sativ
37	6	15.0	447	1	EF1A_SOYBN	P25698 glycine max
38	6	15.0	447	1	EF1A_VICFA	Q24534 vicia faba
39	6	15.0	447	1	EF1A_WHEAT	Q03033 triticum ae
40	6	15.0	449	1	EF11_DAVCA	P29521 daucus caro
41	6	15.0	451	1	MYBH_DICDI	P34127 dictyosceli
42	6	15.0	473	1	C742_PARAR	Q40778 parthenium
43	6	15.0	498	1	YAOI_SCHPO	Q10097 schizosacch
44	6	15.0	518	1	CP74_ARATH	Q26242 arabidopsis
45	6	15.0	524	1	GUAA_YEAST	P38625 saccharomyc

ALIGNMENTS

RESULT 1	EF1A_MANES	STANDARD;	PRT;	449 AA.
ID	EF1A_MANES			
AC	O49169;			
DT	15-DEC-1998 (Rel. 37, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	15-DEC-1998 (Rel. 37, Last annotation update)			
DE	Elongation factor 1-alpha (EF-1-alpha).			
GN	EF1.			
OS	Manihot esculenta (Cassava) (Manioc).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;			
OC	eurosidia I; Malpighiales; Euphorbiaceae; Manihot.			
OX	NCBI_TaxID=3983;			
ON	[1]			
RP	SEQUENCE FROM N. A.			
RC	STRAIN=CV. MIRA 534;			
RA	Sunandono S., Hughes J., Brown K., Sirju-Charan G., Hughes M.A.;			
RT	"An elongation factor 1-alpha gene from cassava (Manihot esculenta			
RL	Crantz.)."			
CC	(In) Plant Gene Register PGR98-052.			
CC	(-) FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF			
CC	AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN			
CC	BIOSYNTHESIS.			
CC	(-) SUBCELLULAR LOCATION: Cytoplasmic.			
CC	(-) SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.			
CC	EF-TU/EF-1A SUBFAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL; AF041463; AAC39447.1; -			
DR	HSSP; P07157; IAIIP.			
DR	InterPro; IPR004539; EF1_alpha.			
DR	InterPro; IPR000795; EF_Grpbind.			
DR	InterPro; IPR004160; EFTU_Cterm.			
DR	InterPro; IPR004161; EFTU_D2.			
DR	Pfam; PF00009; GTP_EFTU_1.			
DR	Pfam; PF03144; GTP_EFTU_D2; 1.			
DR	Pfam; PF03143; GTP_EFTU_D3; 1.			
DR	PRINTS; PR00315; ELONGATNFCT.			
DR	TIGRPFAMs; TIGR00483; EF-1_alpha; 1.			
DR	PROSITE; PS00301; EFACOR_GTP; 1.			
KW	Elongation factor; Protein biosynthesis; GTP-binding.			
FT	NP BIND 14			
FT	NP BIND 21			
FT	NP BIND 91			
FT	NP BIND 95			
FT	NP BIND 153			
SO	SEQUENCE 449 AA; 49371 MW; 779E3858DBE104 CRC64;			

Query Match 17.5%; Score 7; DB 1; Length 449;
Best Local Similarity 100.0%; Pred. No. 3.2;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 VEKDDPS 10
Db 429 VEKDDPS 435

RESULT 2

SPS2_CRAPL STANDARD; PRT; 1081 AA.
ID_SPS2_CRAPL 004933;
AC 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 15-DEC-1998 (Rel. 37, Last annotation update).
DE Sucrose-phosphate synthase 2 (EC 2.4.1.14) (UDP-glucose-fructose-phosphate glucosyltransferase 2).
GN SPS2.
OS Craterostigma plantaginum.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; Lamiales; Lamiales; incertae sedis; Toreniaeae; OC Craterostigma.
OC Craterostigma.
OX NCBI_TaxID=4153;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97451773; PubMed=9306694;
RA Ingram J., Chander J.W., Gallagher L., Salamini F., Bartels D.;
RT Analysis of cDNA clones encoding sucrose-phosphate synthase in relation to sugar interconversions associated with dehydration in the resurrection plant Craterostigma plantaginum Hochst.;
RL Plant Physiol. 115:113-121(1997).
CC -1- FUNCTION: INVOLVED IN THE REGULATION OF CARBON PARTITIONING IN THE LEAVES OF PLANTS. MAY REGULATE THE SYNTHESIS OF SUCROSE AND THEREFORE PLAY A MAJOR ROLE AS A LIMITING FACTOR IN THE EXPORT OF PHOTOSYNTHETIC PRODUCTS OUT OF THE LEAF.
CC -1- CATALYTIC ACTIVITY: UDP-glucose + D-fructose 6-phosphate = UDP + sucrose 6-phosphate.
CC -1- ENZYME REGULATION: ACTIVITY REGULATED BY PHOSPHORYLATION AND MODERATED BY CONCENTRATION OF METABOLITES AND LIGHT.
CC -1- PATHWAY: Sucrose synthesis.
CC -1- SUBUNIT: Homodimer or homotetramer (By similarity).
CC -1- PTM: PHOSPHORYLATED. HOWEVER, PHOSPHORYLATION IS NOT ESSENTIAL FOR ENZYME FUNCTION (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
CC -----
CC EMBL, Y11795; CAA72491.1; -.
DR PIR: T09837; T09837.
DR InterPro: IPR001296; Glyco_transf_1.
DR Pfam: PF00534; Glyco_transf_1.1.
KW Transferase; Glycosyltransferase; Phosphorylation; Multigene family.
KW DOMAIN 248 POLY-SER.
FT DOMAIN 256 264 POLY-GLU.
FT DOMAIN 787 790 POLY-ARG.
SQ SEQUENCE 1081 AA; 120933 MW; DD142DC2F1A72900 CRC64;

Query Match 17.5%; Score 7; DB 1; Length 1081;
Best Local Similarity 100.0%; Pred. No. 6.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 KTISKLM 36
Db 903 KTISKLM 909

RESULT 3

ACYM_RAT STANDARD; PRT; 96 AA.
ID_ACYM_RAT P35745;
AC 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Acylphosphatase, muscle type isozyme (EC 3.6.1.7) (Acylphosphate phosphohydrolyase).
GN ACYP2 OR ACYP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE.
RC TISSUE=Skeletal muscle;
RX MEDLINE=91273765; PubMed=1647162;
RA Berti A., Tremori E., Pazzagli L., Degl'Innocenti D.,
RT "Rat muscle acylphosphatase: purification, amino sequence, and immunological characterization."
RL J. Protein Chem. 10:91-102(1991).
CC -1- FUNCTION: ITS PHYSIOLOGICAL ROLE IS NOT YET CLEAR.
CC -1- CATALYTIC ACTIVITY: An acyl phosphate + H(2)O = a fatty acid anion + phosphate.
CC -1- SIMILARITY: BELONGS TO THE ACYLPHOSPHATASE FAMILY.
CC PIR: A61427; A61427.
DR HSSP: P00818; IAPS.
DR InterPro: IPR001792; Acylphosphatase.
DR Pfam: PF00708; Acylphosphatase; 1.
DR PRINTS: PR00112; ACYLPHPHATSE.
DR PRODOM: P0001884; Acylphosphatase; 1.
DR PROSITE: PS00150; ACYLPHOSPHATASE_1; 1.
DR PROSITE: PS00151; ACYLPHOSPHATASE_2; 1.
KW Hydrolyase; Acetylation; Multigene family.
FT MOD RES 1 1 ACETYLATION.
FT ACT SITE 21 21 ACETYLATION.
FT ACT SITE 39 39 POTENTIAL.
SQ SEQUENCE 96 AA; 10732 MW; 726D5B9B6820D4B CRC64;

Query Match 15.0%; Score 6; DB 1; Length 96;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 KTISKL 35
Db 82 KTISKL 87

RESULT 4

ACYM_MOUSE STANDARD; PRT; 97 AA.
ID_ACYM_MOUSE P56375; Q8B0X2;
AC 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Acylphosphatase, muscle type isozyme (EC 3.6.1.7) (Acylphosphate phosphohydrolyase).
GN ACYP2 OR ACYP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Marra M., Hillier L., Allen M., Bowles M., Dietrich N., Dubuque T., Geisel S., Kucaba T., Lacy M., Le M., Martin J., Morris M., Schellenberg K., Steppe M., Tan F., Underwood K., Moore B., Theising B., Wylie F., Lennon G., Soares B., Wilson R., Waterston R.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina, and Tongue;

RA PubMed12466851;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nishikido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gotojori T.,
 RA Balderelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matveeva H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Bruscia V., Chochia C., Corbani L.E., Cousins S.,
 RA Dalia E., Dregant T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincin S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Konegata A., Kurochkin I.V., Lee Y., Lennard B., Lyons P.A.,
 RA Maglott D.R., Maltas L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pereira G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Raveisi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sadelin A., Schneider C., Sempke C.A., Setou M., Shindake K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita K.,
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wyszewski-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carinci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yaenunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.,
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs."
 RL Nature 420:563-573(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Straube R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shermen C.M., Schuler G.D.,
 RA Altshul S.F., Zeeberg B., Buelow K.H., Scheffer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Martushina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Schetz T.E.,
 RA Brownstein M.J., Ueda T.B., Toshitsuki S., Carrinci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwen P.C., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliy S.W.,
 RA Villalón D.K., Wuzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- CATALYTIC ACTIVITY: An acyl phosphate + H(2)O = a fatty acid anion
 CC + phosphate.
 CC -1- SIMILARITY: BELONGS TO THE ACYLPHOSPHATASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AA285985; -; NOT_ANNOTATED_CDS.
 DR EMBL; AA120674; -; NOT_ANNOTATED_CDS.
 DR EMBL; AK046238; BAC32649.1; ALT_INIT.
 DR EMBL; AK009134; BAC26095.2; ALT_INIT.
 DR EMBL; BC027642; AAH27642.1; -.
 DR HSSP; P00818; IAPS.
 DR MGD; MGI:1922822; Acyp2.

DR InterPro; IPR001792; Acylphosphatase.
 DR Pfam; PF00708; Acylphosphatase; 1.
 DR PRINTS; PR00112; ACYLPHPTASE.
 DR ProDom; PD001884; Acylphosphatase; 1.
 DR PROSITE; PS00150; ACYLPHOSPHATASE_1; 1.
 DR PROSITE; PS00151; ACYLPHOSPHATASE_2; 1.
 KW Hydroxylase; Multigene family.
 FT ACT_SITE 22 22 POTENTIAL.
 FT ACT_SITE 40 40 POTENTIAL.
 FT CONFLICT 13 13 F -> L (IN REF. 2; BAC32649).
 FT CONFLICT 54 54 E -> K (IN REF. 2; BAC32649).
 SQ SEQUENCE 97 AA; 10895 MW; 0F3510AFC2AA15B CRC64;

Query Match 15.0%; Score 6; DB 1; Length 97;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 30 KTISKL 35

Db 83 KTISKL 88

RESULT 5
 ID ACYM_BOVIN STANDARD; PRT; 98 AA.
 AC P07033;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Acylphosphatase, muscle type isozyme (EC 3.6.1.7) (Acylphosphate
 DE phosphohydrolyase).
 GN ACYP2 OR ACYP.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Muscle;
 RX MEDLINE=86223180; PubMed=3011706;
 RA Camici G., Manao G., Modesti A., Stefani M., Bertl A., Capugni G.,
 RA Rampoli G.;
 RT "The complete amino acid sequence of bovine skeletal muscle
 RT acylphosphatase."
 RT Ital. J. Biochem. 35:1-15(1986).

CC -1- FUNCTION: ITS PHYSIOLOGICAL ROLE IS NOT YET CLEAR.
 CC -1- CATALYTIC ACTIVITY: An acyl phosphate + H(2)O = a fatty acid anion
 CC + phosphate.
 CC -1- SIMILARITY: BELONGS TO THE ACYLPHOSPHATASE FAMILY.
 CC PIR; A29578; QPBO.
 DR HSSP; P00818; IAPS.
 DR InterPro; IPR001792; Acylphosphatase.
 DR Pfam; PF00708; Acylphosphatase; 1.
 DR PRINTS; PR00112; ACYLPHPTASE.
 DR ProDom; PD001884; Acylphosphatase; 1.
 DR PROSITE; PS00150; ACYLPHOSPHATASE_1; 1.
 DR PROSITE; PS00151; ACYLPHOSPHATASE_2; 1.
 KW Hydroxylase; Multigene family.
 FT ACT_SITE 23 23 ACETYLATION.
 FT ACT_SITE 41 41 POTENTIAL.
 SQ SEQUENCE 98 AA; 11046 MW; 4B0262FFAD3E06 CRC64;

Query Match 15.0%; Score 6; DB 1; Length 98;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 30 KTISKL 35

Db 84 KTISKL 89

RESULT 6

ACTM_HORSE	STANDARD;	PRT;	98 AA.
ID ACYM_HORSE			
AC P00818;			
DT 21-JUL-1986 (Rel. 01, Created)			
DT 21-JUL-1986 (Rel. 01, Last sequence update)			
DT 28-FEB-2003 (Rel. 41, Last annotation update)			
DE Acylphosphatase, muscle type isozyme (EC 3.6.1.7) (Acylphosphate phosphohydrolyase).			
GN ACYP2 OR ACYP.			
OS Equus caballus (Horse).			
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.			
OX NCBI_Taxid=9796;			
RN (1)			
RP SEQUENCE.			
RC TISSUE=Muscle;			
RX MEDLINE=80227852; PubMed=6248536;			
RA Capugli G., Manao G., Camici G., Rampont G.;			
RT "The complete amino acid sequence of horse muscle acylphosphatase.";			
RL J. Biol. Chem. 255:6868-6874(1980).			
RN (2)			
RP STRUCTURE BY NMR OF ALPHA HELICES.			
RX MEDLINE=89178637; PubMed=2538623;			
RA Saudak V., Atkinson R.A., Williams R.J.P., Rampont G.;			
RT "Identification and description of alpha-helical regions in horse muscle acylphosphatase by 1H nuclear magnetic resonance spectroscopy.";			
RL J. Mol. Biol. 205:229-239(1989).			
RN (3)			
RP STRUCTURE BY NMR OF BETA-STRUCTURES.			
RX MEDLINE=89329032; PubMed=2537076;			
RA Saudak V., Wormald M.R., Williams R.J.P., Boyd J., Stefani M., Rampont G.;			
RT "Identification and description of beta-structure in horse muscle acylphosphatase by nuclear magnetic resonance spectroscopy.";			
RL J. Mol. Biol. 207:405-415(1989).			
RN (4)			
RP STRUCTURE BY NMR.			
RX MEDLINE=92219262; PubMed=1313885;			
RA Pastore A., Saudak V., Rampont G., Williams R.J.P.;			
RT "Three-dimensional structure of acylphosphatase. Refinement and structure analysis.";			
RL J. Mol. Biol. 224:427-440(1992).			
CC -1- FUNCTION: ITS PHYSIOLOGICAL ROLE IS NOT YET CLEAR.			
CC -1- CATALYTIC ACTIVITY: An acyl phosphate + H(2)O = a fatty acid anion + phosphate.			
CC -1- SIMILARITY: BELONGS TO THE ACYLPHOSPHATASE FAMILY.			
CC PIR: A01015; OPHO.			
DR PDB: 1APS; 15-JUL-92.			
DR InterPro: IPR001792; Acylphosphatase.			
DR Pfam: PF00708; Acylphosphatase; 1.			
DR PRINTS: PR00112; ACYLPHPTASE.			
DR PRODOM: PD001884; Acylphosphatase; 1.			
DR PROSITE: PS00150; ACYLPHOSPHATASE_1; 1.			
DR PROSITE: PS00151; ACYLPHOSPHATASE_2; 1.			
KW Hydrolyase; Acetylation; Glutathionylation; Multigene family; 3D-structure.			
KW 3D-structure.			
FT ACT SITE	23	POTENTIAL.	
FT ACT SITE	41	POTENTIAL.	
FT MOD RES	1	ACETYLATION.	
FT MOD RES	21	GLUTATHIONYLATION.	
FT TUR	2		
FT TUR	3		
FT STRAND	6		
FT TUR	16		
FT TUR	17		
FT TUR	23		
FT HELIX	25		
FT TUR	33		
FT TUR	34		
FT STRAND	36		
FT TUR	40		
FT STRAND	43		
FT STRAND	46		
FT HELIX	54		
FT STRAND	55		
FT STRAND	78		

FT STRAND 94 98

SEQ SEQUENCE 98 AA; 11016 MW; 04B4CEB1B03C3E00 CRC64;

Query Match

Best Local Similarity 100.0%; Score 6; DB 1; Length 98;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 30 KTISKL 35

DB 84 KTISKL 89

RESULT 7

ACTM_HUMAN	STANDARD;	PRT;	98 AA.
ID ACYM_HUMAN			
AC P14621;			
DT 01-APR-1990 (Rel. 14, Created)			
DT 01-APR-1990 (Rel. 14, Last sequence update)			
DT 15-SEP-2003 (Rel. 42, Last annotation update)			
DE Acylphosphatase, muscle type isozyme (EC 3.6.1.7) (Acylphosphate phosphohydrolyase).			
GN ACYP2 OR ACYP.			
OS Homo sapiens (Human).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX NCBI_Taxid=9606;			
RN (1)			
RP SEQUENCE.			
RC TISSUE=Muscle;			
RX MEDLINE=86174259; PubMed=6100723;			
RA Manao G., Camici G., Modesti A., Liguori G., Berti A., Sefiani M., Capugli G., Rampont G.;			
RT "Human skeletal muscle acylphosphatase: the primary structure.";			
RL Mol. Biol. Med. 2:369-378(1984).			
RN (2)			
RP SEQUENCE FROM N.A.			
RC TISSUE=Testis;			
RA Hillier L., Allen M., Bowles M., Dubuque T., Geisel S., Jost S., Kucaba T., Lacy M., Le M., Lennon G., Marra M., Martin J., Moore B., Schellenberg K., Steptoe M., Tan F., Theising B., White Y., Wylie T., Watson R., Wilson R.;			
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.			
RN (3)			
RP SEQUENCE FROM N.A.			
RC TISSUE=Ovary;			
RX MEDLINE=22388257; PubMed=12477932;			
RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G., Altschul S.F., Zeeberg B., Buelow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E., Brownstein M.J., Usdin T.B., Toshlyuk S., Carrinoni P., Prange C., Rane S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Boask S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hilyk S.W., Villalón D.K., Wuzny K.C., Sodergren E.J., Lu X., Gibbs R.A., Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smallie D.E., Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;			
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";			
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).			
RN (4)			
RP SEQUENCE OF 21-98 FROM N.A.			
RC TISSUE=Heart;			
RX MEDLINE=96033055; PubMed=7487897;			
RA Chiarugi P., Raugi G., Marzocchi R., Fiaschi T., Ciccarelli C., Berti A., Rampont G.;			
RT "Differential modulation of expression of the two acylphosphatase			

```

RT isoenzymes by thyroid hormone.";
RL Biochem. J. 311:567-573(1995).
CC -1- FUNCTION: ITS PHYSIOLOGICAL ROLE IS NOT YET CLEAR.
CC -1- CATALYTIC ACTIVITY: An acyl phosphate + H(2)O = a fatty acid anion
CC + phosphate.
CC -1- SIMILARITY: BELONGS TO THE ACYLPHOSPHATASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AA405029; -; NOT ANNOTATED_CDS.
DR EMBL; BC012290; AA012250.1; -.
DR EMBL; X84195; CA58988.1; -.
DR PIR; S59138; S52327.
DR HSSP; P00818; IAPS.
DR Genew; HGNC:180; ACYP2.
DR MIM; 102595; -.
DR GO; GO:0003998; F:acylphosphatase activity; TAS.
DR GO; GO:0006796; P:phosphate metabolism; TAS.
DR InterPro; IPR001792; Acylphosphatase.
DR Pfam; PF00708; Acylphosphatase; 1.
DR PRINTS; PR00112; ACYLPHPTASE.
DR PRODOM; PD001884; ACYLPHOSPHATASE_1; 1.
DR PROSITE; PS00150; ACYLPHOSPHATASE_2; 1.
DR PROSITE; PS00151; ACYLPHOSPHATASE_2; 1.
DR HydroLase; Acetylation; Multigene family.
GN INIT MET 0 0
FT MOD RES 1 1 ACETYLATION.
FT ACT_SITE 23 23 POTENTIAL.
FT ACT_SITE 41 41 POTENTIAL.
SQ SEQUENCE 98 AA; 11008 MW; 3877201DCBFB0E30 CRC64;

Query Match 15.0%; Score 6; DB 1; Length 98;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 30 KTISKL 35
DB 84 KTISKL 89

RESULT 8
ACYM_PIG STANDARD; PRT; 98 AA.
ID ACYM_PIG
AC P00819;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Acylphosphatase, muscle type isozyme (EC 3.6.1.7) (Acylphosphate
DE phosphohydrolyase).
GN ACYP2 OR ACYP.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9623;
RN [1]
RP SEQUENCE.
RC TISSUE=Skeletal muscle;
RX MEDLINE=85289095; PubMed=2993259;
RA Mizuno Y., Yamazaki M., Takasawa T., Kizaki T., Shiokawa H.;
RT "Amino acid sequence of acylphosphatase from porcine skeletal
RT muscle.";
RL J. Biochem. 97:1135-1142(1985).
CC -1- FUNCTION: ITS PHYSIOLOGICAL ROLE IS NOT YET CLEAR.
CC -1- CATALYTIC ACTIVITY: An acyl phosphate + H(2)O = a fatty acid anion
CC + phosphate.
CC -1- SIMILARITY: BELONGS TO THE ACYLPHOSPHATASE FAMILY.
CC PIR; A01016; QPFG.

```

```

DR HSSP; P00818; IAPS.
DR InterPro; IPR001792; Acylphosphatase.
DR Pfam; PF00708; Acylphosphatase; 1.
DR PRINTS; PR00112; ACYLPHPTASE.
DR PRODOM; PD001884; Acylphosphatase; 1.
DR PROSITE; PS00150; ACYLPHOSPHATASE_1; 1.
DR PROSITE; PS00151; ACYLPHOSPHATASE_2; 1.
DR HydroLase; Acetylation; Multigene family.
FT MOD RES 1 1 ACETYLATION.
FT ACT_SITE 23 23 POTENTIAL.
FT ACT_SITE 41 41 POTENTIAL.
SQ SEQUENCE 98 AA; 11074 MW; 9F35DC61A3D20F1 CRC64;

Query Match 15.0%; Score 6; DB 1; Length 98;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 30 KTISKL 35
DB 84 KTISKL 89

RESULT 9
ACYM_RABIT STANDARD; PRT; 98 AA.
ID ACYM_RABIT
AC P00820;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Acylphosphatase, muscle type isozyme (EC 3.6.1.7) (Acylphosphate
DE phosphohydrolyase).
GN ACYP2 OR ACYP.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE.
RC TISSUE=Skeletal muscle;
RX MEDLINE=85289097; PubMed=2993260;
RA Kizaki T., Takasawa T., Mizuno Y., Shiokawa H.;
RT "Amino acid sequence of acylphosphatase from rabbit skeletal muscle.";
RL J. Biochem. 97:1155-1161(1985).
RN [2]
RP SEQUENCE.
RC TISSUE=Skeletal muscle;
RX MEDLINE=85305733; PubMed=2994566;
RA Manao G., Camici G., Capugni G., Stefani M., Liguri G., Bertl A.,
RA Rampoli G.;
RT "Rabbit skeletal muscle acylphosphatase: the amino acid sequence of
RT form Ral.";
RL Arch. Biochem. Biophys. 241:418-423(1985).
CC -1- FUNCTION: ITS PHYSIOLOGICAL ROLE IS NOT YET CLEAR.
CC -1- CATALYTIC ACTIVITY: An acyl phosphate + H(2)O = a fatty acid anion
CC + phosphate.
CC -1- SUBUNIT: Monomer (RA1) and homodimer (RA3); disulfide linked.
CC -1- SIMILARITY: BELONGS TO THE ACYLPHOSPHATASE FAMILY.
CC PIR; A91988; QPRB.
DR HSSP; P00818; IAPS.
DR InterPro; IPR001792; Acylphosphatase.
DR Pfam; PF00708; Acylphosphatase; 1.
DR PRINTS; PR00112; ACYLPHPTASE.
DR PRODOM; PD001884; ACYLPHOSPHATASE_1; 1.
DR PROSITE; PS00150; ACYLPHOSPHATASE_2; 1.
DR PROSITE; PS00151; ACYLPHOSPHATASE_2; 1.
DR HydroLase; Acetylation; Multigene family.
FT MOD RES 1 1 ACETYLATION.
FT DISULFID 21 21 INTERCHAIN (PARTIAL).
FT MOD RES 21 21 GLUTATHIONYLATION (PARTIAL).
FT ACT_SITE 23 23 POTENTIAL.
FT ACT_SITE 41 41 POTENTIAL.
SQ SEQUENCE 98 AA; 10861 MW; 8B10CA8568683FB0 CRC64;

```

Query Match 15.0%; Score 6; DB 1; Length 98;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 30 KTISKL 35
 |||||
 Db 84 KTISKL 89

RESULT 10

MSMB_PIG STANDARD; PRT; 111 AA.
 ID MSMB_PIG 002826;
 AC 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Beta-microseminoprotein precursor (Prostate secreted seminal plasma protein) (Prostate secretory protein PSP94) (PSP-94) (TSS07).
 GN MSMB OR PSP94.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96117078; PubMed=8562060;
 RA Tanaka T., Itahana K., Andon N., Takeya T., Sato E.;
 RT "Expression of prostatic secretory protein (PSP)-like protein in
 RT porcine corpus luteum: isolation and characterization of a new gene
 RT encoding PSP94-like protein.";
 RL Mol. Reprod. Dev. 42:149-156(1995).
 RL [2]
 RP SEQUENCE OF 21-111.
 RC TISSUE=Seminal plasma;
 RX MEDLINE=94161559; PubMed=8117114;
 RA Ferlund P., Granberg L.B., Roepstorff P.;
 RT "Amino acid sequence of beta-microseminoprotein from porcine seminal
 RT plasma.";
 RL Arch. Biochem. Biophys. 309:70-76(1994).
 RL -1- FUNCTION: Inhibits the secretion of FSH by pituitary cells (By
 RL similarity).
 CC -1- SUBCELLULAR LOCATION: Secreted. Sperm surface (By similarity).
 CC -1- TISSUE SPECIFICITY: CORPORA LUTEA, MOSTLY IN THE LUTEAL CELLS
 CC SURROUNDING BLOOD VESSELS.
 CC -1- SIMILARITY: BELONGS TO THE BETA-MICROSEMINOPROTEIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; S80724; AAB50711.1; -
 KW Signal; Pyrolydione carboxylic acid.
 FT SIGNAL 1 20
 FT CHAIN 21 111
 FT DISULFID 22 35 BETA-MICROSEMINOPROTEIN.
 FT DISULFID 54 90 BY SIMILARITY.
 FT DISULFID 57 66 OR 67 (BY SIMILARITY).
 FT DISULFID 59 67 OR 66 (BY SIMILARITY).
 FT DISULFID 81 104 BY SIMILARITY.
 FT MOD_RES 21 21 PYROLYDIONE CARBOXYLIC ACID.
 SQ SEQUENCE 111 AA; 12246 MW; ABR33A067A23765E CRC64;
 Query Match 15.0%; Score 6; DB 1; Length 111;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 11

MSPA_SAGE STANDARD; PRT; 112 AA.
 ID MSPA_SAGE 097936;
 AC 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Beta-microseminoprotein A1 precursor (msp-A1) (Fragment).
 GN MSPA.
 OS Saguinus oedipus (Cotton-top tamarin).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae; Saguinus.
 OX NCBI_TaxID=9490;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99421644; PubMed=10491085;
 RA Maekinen M., Valtanen-Andre C., Lundwall A.;
 RT "New world, but not old world, monkeys carry several genes encoding
 RT beta-microseminoprotein.";
 RL Eur. J. Biochem. 264:407-414(1999).
 RL -1- FUNCTION: Inhibits the secretion of FSH by pituitary cells (By
 RL similarity).
 CC -1- SUBCELLULAR LOCATION: Secreted. Sperm surface (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE BETA-MICROSEMINOPROTEIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AJ010158; CAB38124.1; -
 DR EMBL; AJ010159; CAB38124.1; JOINED.
 DR GO; GO:0005576; C:extracellular; TMS.
 DR InterPro; IPR002400; GF_cysknoc.
 KW Signal.
 FT NON_TER 1 1
 FT SIGNAL <1 19
 FT CHAIN 20 112 BETA-MICROSEMINOPROTEIN A1.
 FT DISULFID 21 37 BY SIMILARITY.
 FT DISULFID 56 92 BY SIMILARITY.
 FT DISULFID 59 68 OR 69 (BY SIMILARITY).
 FT DISULFID 61 69 OR 68 (BY SIMILARITY).
 FT DISULFID 83 106 BY SIMILARITY.
 SQ SEQUENCE 112 AA; 12631 MW; DDA05BFC15CD78B CRC64;
 Query Match 15.0%; Score 6; DB 1; Length 112;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 VEKDP 9
 |||||
 Db 97 VEKDP 102

RESULT 12

MSMB_HUMAN STANDARD; PRT; 114 AA.
 ID MSMB_HUMAN 097936;
 AC 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Beta-microseminoprotein precursor (Prostate secreted seminal plasma
 DE protein) (Prostate secretory protein PSP94) (PSP-94) (Seminal plasma
 DE beta-inhibin) (Immunoglobulin binding factor) (IGBF) (PM44).
 GN MSMB OR PRSP.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

CX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87161231; PubMed=3829888;
 RA Mbikay M., Nolet S., Fournier S., Benjannet S., Chapelaine P.,
 RA Paradis G., Dube J.Y., Tremblay R., Lazure C., Seidah N.G.,
 RA Chretien M.;
 RT "Molecular cloning and sequence of the cDNA for a 94-amino-acid
 RT seminal plasma protein secreted by the human prostate.";
 RL DNA 6:23-29(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91274357; PubMed=2054385;
 RA Nolet S., Mbikay M., Chretien M.;
 RT "Prostatic secretory protein PSP94: gene organization and promoter
 RT sequence in rhesus monkey and human";
 RL Biochim. Biophys. Acta 1089:247-249(1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90211299; PubMed=2322265;
 RA Green C.B., Liu W.Y., Kwok S.C.M.;
 RT "Cloning and nucleotide sequence analysis of the human beta-
 RT microseminoprotein gene.";
 RL Biochem. Biophys. Res. Commun. 167:1184-1190(1990).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90073664; PubMed=2590204;
 RA Ulysbach M., Lindstroem C., Weiber H., Abrahamson P.-A., Lilja H.,
 RA Lundwall A.;
 RT "Molecular cloning of a small prostate protein, known as beta-
 RT microseminoprotein, PSP94 or beta-inhibin, and demonstration of
 RT transcripts in non-genital tissues.";
 RL Biochem. Biophys. Res. Commun. 164:1310-1315(1989).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Prostate;
 RX MEDLINE=94115955; PubMed=7506990;
 RA Liu A.Y., Bradner R.C., Vessella R.L.;
 RT "Decreased expression of prostatic secretory protein PSP94 in
 RT prostate cancer";
 RL Cancer Lett. 74:91-99(1993).
 RN [6]
 RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.
 RC TISSUE=Prostate;
 RX MEDLINE=96032566; PubMed=7566962;
 RA Xuan J.W., Chin J.L., Guo Y., Chambers A.F., Finkelman M.A.,
 RA Clarke M.W.;
 RT "Alternative splicing of PSP94 (prostatic secretory protein of 94
 RT amino acids) mRNA in prostate tissue.";
 RL Oncogene 11:1041-1047(1995).
 RN [7]
 RP SEQUENCE FROM N.A.
 RA Bajjal-Gupta M., Clarke M.W.;
 RT "Prostate specific protein (PSP94) expression in a human endometrial
 RT cell line (KLE)";
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
 RN [8]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Prostate;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Ditschenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ueda T.B., Yoshizuki S., Carrinci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McWen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shvachenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [9]
 RP SEQUENCE OF 2-72 FROM N.A.
 RX MEDLINE=99421644; PubMed=10491085;
 RA Maekinen M., Valtanen-Andre C., Lundwall A.;
 RT "New world, but not old world, monkeys carry several genes encoding
 RT beta-microseminoprotein";
 RL Eur. J. Biochem. 264:407-414(1999).
 RN [10]
 RP SEQUENCE OF 21-113.
 RX MEDLINE=8519974; PubMed=3995056;
 RA Akiyama K., Yoshioke Y., Schmid K., Offner G.D., Troxler R.F.,
 RA Tsuda R., Hara M.;
 RT "The amino acid sequence of human beta-microseminoprotein";
 RL Biochim. Biophys. Acta 829:288-294(1985).
 RN [11]
 RP SEQUENCE OF 21-114.
 RX MEDLINE=85004133; PubMed=6434350;
 RA Seidah N.G., Arpatti N.J., Rochemont J., Sheeh A.R., Chretien M.;
 RT "Complete amino acid sequence of human seminal plasma beta-inhibin.
 RT Prediction of post Glu-Arg cleavage as a maturation site.";
 RL FEBS Lett. 175:349-355(1984).
 RN [12]
 RP SEQUENCE OF 21-50 AND 113-114.
 RX MEDLINE=92028964; PubMed=1930232;
 RA Liang Z.G., Kamada M., Koide S.S.;
 RT "Structural identity of immunoglobulin binding factor and prostatic
 RT secretory protein of human seminal plasma";
 RL Biochem. Biophys. Res. Commun. 180:356-359(1991).
 RN [13]
 RP SEQUENCE OF 21-41, AND TISSUE SPECIFICITY.
 RC TISSUE=Semen;
 RX MEDLINE=95401076; PubMed=7671139;
 RA Ohkubo I., Tada T., Ochiai Y., Ueyama H., Himoto T., Saeki M.;
 RT "Human seminal plasma beta-microseminoprotein: its purification,
 RT characterization, and immunohistochemical localization";
 RL Int. J. Biochem. Cell Biol. 27:603-611(1995).
 RN [14]
 RP SEQUENCE OF 21-32.
 RX MEDLINE=21648993; PubMed=11788998;
 RA Ghafoori B., Stahlbow B., Tagesson C., Lindahl M.;
 RT "Newly identified proteins in human nasal lavage fluid from
 RT non-smokers and smokers using two-dimensional gel electrophoresis and
 RT peptide mass fingerprinting";
 RL Proteomics 2:112-120(2002).
 RN [15]
 RP FUNCTION: Inhibits the secretion of FSH by pituitary cells.
 CC -1- SUBCELLULAR LOCATION: Secreted. Sperm surface.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=PSP94;
 CC IsoId=P08118-1; Sequence=Displayed;
 CC Name=PSP57;
 CC IsoId=P08118-2; Sequence=VSP_003275, VSP_003276;
 CC -1- TISSUE SPECIFICITY: Strongly expressed in prostate, liver, kidney,
 CC breast and penis. Also expressed in pancreas, esophagus, stomach,
 CC duodenum, colon, trachea, lung, salivary glands and fallopian
 CC tube. PSP94 is expressed in lung and breast, whereas PSP57 is
 CC found in kidney and bladder.
 CC -1- MISCELLANEOUS: SPECIFIC RECEPTORS FOR THIS PROTEIN ARE FOUND ON
 CC SPERMATOZOEA AND IN THE PROSTATE.
 CC -1- SIMILARITY: BELONGS TO THE BETA-MICROSEMINOPROTEIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

```

CC FORM A SWITCH COMPLEX THAT IS PROPOSED TO BE LOCATED AT THE BASE
CC OF THE BASAL BODY. THIS COMPLEX INTERACTS WITH THE CHEY AND CHEZ
CC CHEMOTAXIS PROTEINS, IN ADDITION TO CONTACTING COMPONENTS OF THE
CC MOTOR THAT DETERMINE THE DIRECTION OF FLAGELLAR ROTATION (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Membrane-associated (By similarity).
CC -1- SIMILARITY: BELONGS TO THE FLIN/MOPA/SPAO FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; AE014083; AAM67644.1; -.
CC DR InterPro; IPR001172; Flagellar_Flin.
CC DR InterPro; IPR001543; Spoa.
CC DR Pfam; PF01052; Spoa; 1.
CC DR PRINTS; PR00956; FLAGMOTORFLIN.
CC DR ProDom; PD001777; Spoa; 1.
CC KW Chemotaxis; Flagella; Flagellar rotation; Membrane; Complete proteome.
CC SO SEQUENCE 133 AA; 15403 MW; CB00B5E2AC8A3286 CRC64;
OY 3 KVEKED 8
Db 3 KVEKED 8
RESULT 14
PAP_ECOLI STANDARD; PRT; 173 AA.
AC P08407;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Flmbrial protein pape precursor.
DE PAPE.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
CX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RP STRAIN=J96;
RX MEDLINE=86149403; Pubmed=2869489;
RA Lindberg F., Lund B., Normark S.;
RT "Gene products specifying adhesion of uropathogenic Escherichia coli
RT are minor components of pili.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:1891-1895(1986).
RN [2]
RP SEQUENCE FROM N.A.
RP STRAIN=J96;
RX MEDLINE=88169520; Pubmed=2895103;
RA Lund B., Lindberg F., Normark S.;
RT "Structure and antigenic properties of the tip-located P pilus
RT proteins of uropathogenic Escherichia coli.";
RL J. Bacteriol. 170:1887-1894(1988).
RN [3]
RP SEQUENCE FROM N.A.
RP STRAIN=J96;
RX MEDLINE=93028852; Pubmed=1357526;
RA Marklund B.-I., Pennett J.M., Garcia E., Hamers A., Baga M.,
RA Lindberg F., Gaastera W., Normark S.;
RT "Horizontal gene transfer of the Escherichia coli pap and prs pili
RT operons as a mechanism for the development of tissue-specific
RT adhesive properties.";
RL Mol. Microbiol. 6:2225-2242(1992).

```

```

RN [4]
RP REVIEW.
RX MEDLINE=99160448; PubMed=10049807;
RA Hung D.L., Hultgren S.J.;
RT "Plus bigenesis via the chaperrone/usher pathway: an integration of
   structure and function.";
RL J. Struct. Biol. 124:201-220(1998).
CC -1- FUNCTION: Repeated pape subunits make up the thin (2 nm in
   diameter) tip fibrillum of the pilus. Subunits are arranged in a
   open helical conformation. Pili with a defective pape gene have
   low adhesive capacity or none; however, the binding property of
   the whole cell is not affected. Pili are polar filaments radiating
   from the surface of the bacterium to a length of 0.5-1.5
   micrometers and numbering 100-300 per cell, and enable bacteria to
   colonize the epithelium of specific host organs.
CC -1- SUBCELLULAR LOCATION: Secreted; tip of the pilus.
CC -1- MISCELLANEOUS: Strains of E.coli that cause infection of the human
   urinary tract produce pap-pili which are hair-like appendages
   consisting of about 1000 helically arranged subunits of the
   protein pape. These pili mediate binding to disaccharide-
   containing glycolipids present on the epithelial cells which line
   the urinary tract.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
   between the Swiss Institute of Bioinformatics and the EMBL outstation -
   the European Bioinformatics Institute. There are no restrictions on its
   use by non-profit institutions as long as its content is in no way
   modified and this statement is not removed. Usage by and for commercial
   entities requires a license agreement (See http://www.isb-sib.ch/announce/
   or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X61338; CAA43556.1; -
DR EMBL; M13339; AAA24280.1; -
DR EMBL; M20146; AAA24288.1; -
DR EMBL; X61339; CAA43568.1; -
DR PIR; A25134; YOECEP.
DR PDB; 1N12; 1I-DEC-02.
DR InterPro; IPR000259; Fimbrial.
DR InterPro; IPR004086; FimbrialPape.
DR Pfam; PF00419; Fimbrial; 1.
DR PRINTS; PRO1555; FIMBRILPAP.
DR FIMBRIA; Cell adhesion; Signal; 3D-structure.
FT SIGNAL 1 24
FT CHAIN 25 173
SQ SEQUENCE 173 AA; 18569 MM; E27577D09C46A863 CRC64;

```

```

Query Match 15.0%; Score 6; DB 1; Length 173;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy 28 ADKTIIS 33
Db 138 ADKTIIS 143

```

```

RESULT 15
YCR2_BACTK STANDARD; PRT; 252 AA.
AC P21733;
DT 01-MAY-1991 (Rel. 18; Created)
DT 01-MAY-1991 (Rel. 18; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE Hypothetical 29.1 kDa protein in CRVA 5' region (ORF2).
OS Bacillus thuringiensis (subsp. kurstaki).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=29339;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HD-1;
RX MEDLINE=89123178; PubMed=2914879;
RA Widner W.R., Whiteley H.R.;
RT "Two highly related insecticidal crystal proteins of Bacillus
   thuringiensis subsp. kurstaki possess different host range

```

```

RT specificities.";
RL J. Bacteriol. 171:965-974(1989).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
   between the Swiss Institute of Bioinformatics and the EMBL outstation -
   the European Bioinformatics Institute. There are no restrictions on its
   use by non-profit institutions as long as its content is in no way
   modified and this statement is not removed. Usage by and for commercial
   entities requires a license agreement (See http://www.isb-sib.ch/announce/
   or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M23723; AAA83515.1; -
DR PIR; B32053; B32053.
KW Hypothetical protein; Repeat.
FT DOMAIN 68 246
FT REPEAT 68 82
FT REPEAT 83 97
FT REPEAT 98 112
FT REPEAT 113 127
FT REPEAT 128 142
FT REPEAT 143 157
FT REPEAT 158 172
FT REPEAT 173 187
FT REPEAT 188 202
FT REPEAT 203 217
FT REPEAT 218 230
FT REPEAT 231 239
FT REPEAT 240 246
SQ SEQUENCE 252 AA; 29110 MM; 382F3BCB1EFDEC9 CRC64;

```

```

Query Match 15.0%; Score 6; DB 1; Length 252;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy 20 GNKIYT 25
Db 48 GNKIYT 53

```

```

Search completed: July 24, 2003, 12:22:10
Job time : 3.63543 secs

```

THIS PAGE BLANK (5871)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 12:16:28 ; Search time 12.2987 Seconds
(without alignments)
839.285 Million cell updates/sec

Title: US-09-991-681-30

Perfect score: 40
Sequence: 1 SPKVEKKDPSRKKEWENAGNKIYTMADTKTISKLTMEYK 40

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size : 0

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

SPTREMBL_23:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phage:*

10: sp_plant:*

11: sp_proteint:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_virus:*

16: sp_bacteriaph:*

17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	589	4 Q8N4Y4	Q8N4Y4 homo sapien
2	40	100.0	592	4 Q96CH9	Q96CH9 homo sapien
3	40	100.0	1770	4 Q9ULH6	Q9ULH6 homo sapien
4	40	100.0	33	10 Q945D4	Q945D4 castanea sa
5	37	17.5	143	10 P93272	P93272 malus domes
6	37	17.5	182	12 Q91FG9	Q91FG9 chilo ride
7	37	17.5	285	16 Q8DTE8	Q8DTE8 streptococc
8	37	17.5	406	5 Q22192	Q22192 caenorhabdi
9	37	17.5	406	10 Q98233	Q98233 arabidopsis
10	37	17.5	407	10 Q8LEJ9	Q8LEJ9 arabidopsis
11	37	17.5	616	5 Q81787	Q81787 plasmodium
12	37	17.5	779	5 Q01441	Q01441 caenorhabdi
13	37	17.5	913	17 Q980C0	Q980C0 sulfolobus
14	37	17.5	1029	10 Q65580	Q65580 arabidopsis
15	37	17.5	1037	10 Q981P2	Q981P2 arabidopsis
16	37	17.5	1096	10 Q8W556	Q8W556 arabidopsis

17	6	15.0	58	13 Q8WVT2	Q8WVT2 myxine glut
18	6	15.0	73	16 Q92C89	Q92C89 listeria in
19	6	15.0	83	5 Q9VXR1	Q9VXR1 drosophila
20	6	15.0	85	2 Q8KNV8	Q8KNV8 bacillus th
21	6	15.0	94	9 Q9T124	Q9T124 bacterioph
22	6	15.0	97	16 Q9WU04	Q9WU04 staphylococ
23	6	15.0	100	10 Q8S2A4	Q8S2A4 oryza sativ
24	6	15.0	106	11 Q8BQX2	Q8BQX2 mus musculu
25	6	15.0	116	16 Q92PF9	Q92PF9 listeria in
26	6	15.0	117	10 Q9SBR6	Q9SBR6 medicago va
27	6	15.0	118	10 Q8LGF9	Q8LGF9 arabidopsis
28	6	15.0	118	10 Q92B33	Q92B33 listeria in
29	6	15.0	124	16 Q9L0T5	Q9L0T5 streptomyce
30	6	15.0	130	10 Q9SBR9	Q9SBR9 cicor ariet
31	6	15.0	136	9 Q8SDA0	Q8SDA0 pseudomonas
32	6	15.0	137	5 Q964U2	Q964U2 procamburus
33	6	15.0	138	10 Q8S1K0	Q8S1K0 oryza sativ
34	6	15.0	149	17 Q9TWG5	Q9TWG5 sulfolobus
35	6	15.0	151	10 Q8VXW8	Q8VXW8 arabidopsis
36	6	15.0	153	10 Q9XET1	Q9XET1 glycine max
37	6	15.0	162	2 Q9S394	Q9S394 lactobacill
38	6	15.0	164	5 Q81AU2	Q81AU2 plasmodium
39	6	15.0	164	11 Q8K261	Q8K261 mus musculu
40	6	15.0	167	16 Q8RD54	Q8RD54 thermoanaer
41	6	15.0	170	5 Q9U1Q1	Q9U1Q1 caenorhabdi
42	6	15.0	177	5 Q9BJE8	Q9BJE8 plasmodium
43	6	15.0	177	5 Q819A6	Q819A6 plasmodium
44	6	15.0	177	5 Q81995	Q81995 plasmodium
45	6	15.0	177	5 Q810M6	Q810M6 plasmodium

ALIGNMENTS

RESULT 1

ID Q8N4Y4 PRELIMINARY; PRT; 589 AA.

AC Q8N4Y4;

DT 01-OCT-2002 (TREMBlrel. 22, Created)

DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)

DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)

DE Similar to KIAA1244 protein (Fragment).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Colon;

RA Strausberg R.;

RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL, BC033191; AAH33191.1; -.

FT NON TER

FT NON TER

SO SEQUENCE

Query Match 589 AA; 66086 MW; 8041BEA348DE65F7 CRC64;

Best Local Similarity 100.0%; Pred. No. 2, 2e-34; Length 589;

Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPKVEKKDPSRKKEWENAGNKIYTMADTKTISKLTMEYK 40

Db 403 SPKVEKKDPSRKKEWENAGNKIYTMADTKTISKLTMEYK 442

RESULT 2

ID Q96CH9 PRELIMINARY; PRT; 592 AA.

AC Q96CH9;

DT 01-DEC-2001 (TREMBlrel. 19, Created)

DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)

DE Hypothetical protein (Fragment).

OS Homo sapiens (Human).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OK NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC014227; AAH14227.1; -.
KM Hypothetical protein.
FT NON_TER
SQ SEQUENCE 592 AA; 6640 MW; F4A1E807B0DF47B5 CRC64;

Query Match
Best Local Similarity 100.0%; Score 40; DB 4; Length 592;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SPKVEKDPSSRKKEWENAGNKIYTMADKITSKLTETK 40
DB 406 SPKVEKDPSSRKKEWENAGNKIYTMADKITSKLTETK 445

RESULT 3
Q9ULH6 PRELIMINARY; PRT; 1770 AA.
AC Q9ULH6: Q96P46;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE BIG3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OK NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RA Hong W.;
RT "KIAA1244 as a novel distantly related member (BIG3) of the BIG1/Sec7p
subfamily of Arf GEFs."
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF413080; AAL04174.1; -.
DR InterPro; IPR000904; Sec7.
DR SMART; SM00222; Sec7.
KM Hypothetical protein.
SQ SEQUENCE 1770 AA; 195845 MW; 5E996E36A6F92AB4 CRC64;

Query Match
Best Local Similarity 100.0%; Score 40; DB 4; Length 1770;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SPKVEKDPSSRKKEWENAGNKIYTMADKITSKLTETK 40
DB 1584 SPKVEKDPSSRKKEWENAGNKIYTMADKITSKLTETK 1623

RESULT 4
Q945D4 PRELIMINARY; PRT; 33 AA.
AC Q945D4;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Translation elongation factor 1 alpha-like protein (Fragment).
OS Caesanea sativa (Sweet chestnut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; Rosidae;
OC eurosid 1; Fagales; Fagaceae; Castanea.
OK NCBI_TaxID=21020;
RN (1)
RP SEQUENCE FROM N.A.
RA Schafleitner R.; Wilhelm E.;
RT "Wound-responsive genes of chestnut."
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

```

```

DR EMBL; AF417301; AAL15889.1; -.
KM Elongation factor.
FT NON_TER
SQ SEQUENCE 33 AA; 3393 MW; 6C3B7346A29E1AAB CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 10; Length 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 VEKKDPS 10
DB 15 VEKKDPS 21

RESULT 5
P93272 PRELIMINARY; PRT; 143 AA.
ID P93272;
AC P93272;
DT 01-MAY-1997 (TRENBLrel. 03, Created)
DT 01-MAY-1997 (TRENBLrel. 03, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE Elongation factor 1 alpha (Fragment).
GN EF-1ALPHA.
OS Malus domestica (Apple) (Malus sylvestris).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; Rosidae;
OC eurosid 1; Rosales; Rosaceae; Maloideae; Malus.
OK NCBI_TaxID=3750;
RN (1)
RP SEQUENCE FROM N.A.
RA STRAIN=Granny Smith;
RA Dong Y.-H.; Janssen B.J.; Bielecki L.L.; Atkinson R.G.; Morris B.A.;
RA Gardner R.C.;
RT "Isolating and characterizing genes differentially expressed early in
apple fruit development."
RL J. Am. Soc. Hortic. Sci. 122:752-757 (1997).
DR EMBL; U80268; AAC06383.1; -.
DR InterPro; IPR004160; EFTU_Cterm.
DR Pfam; PF03143; GTP_EFTU_D3; 1.
FT NON_TER
SQ SEQUENCE 143 AA; 15625 MW; F047F8ACCC9326DE CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 10; Length 143;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 VEKKDPS 10
DB 125 VEKKDPS 131

RESULT 6
Q91FG9 PRELIMINARY; PRT; 182 AA.
ID Q91FG9;
AC Q91FG9;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE 35SR.
OS Chilo iridescent virus (CIV) (Insect iridescent virus type 6).
OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus.
OK NCBI_TaxID=10488;
RN (1)
RP SEQUENCE FROM N.A.
RA Delius H.; Darai G.; Fluegel R.M.;
RA "DNA analysis of insect iridescent virus 6: evidence for circular
permutation and terminal redundancy."
RL J. Virol. 49:609-614 (1984).
RN (2)
RP SEQUENCE FROM N.A.
RX MEDLINE=86174607; PubMed=3959991;
RA Lorbacher de Ruiz H.; Gelderblom H.; Hofmann W.; Darai G.;
RT "Insect iridescent virus type 6 induced toxic degenerative hepatitis

```

RT in mice.";
 RL Med. Microbiol. Immunol. 175:43-53(1986).
 [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87321126; PubMed=2820141;
 RA Schnitzler P., Soltan J.B., Fischer M., Reissner H., Scholz J.,
 RT "Molecular cloning and physical mapping of the genome of insect
 RL iridescent virus type 6: further evidence for circular permutation of
 the viral genome.";
 RL Virology 160:66-74(1987).
 [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89073752; PubMed=3201750;
 RA Fischer M., Schnitzler P., Delius H., Darai G.;
 RT "Identification and characterization of the repetitive DNA element in
 RL the genome of insect iridescent virus type 6.";
 RL Virology 167:485-496(1988).
 [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92196996; PubMed=1549908;
 RA Handermann M., Schnitzler P., Rosen-Wolfe A., Raab K., Sonntag K.C.,
 RT "Identification and mapping of origins of DNA replication within the
 RL DNA sequences of the genome of insect iridescent virus type 6.";
 RL Virus Genes 6:19-32(1992).
 [6]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93118242; PubMed=1475907;
 RA Sonntag K.C., Darai G.;
 RT "Characterization of the third origin of DNA replication of the genome
 RL of insect iridescent virus type 6.";
 RL Virus Genes 6:333-342(1992).
 [7]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93260401; PubMed=8492091;
 RA Stohwasser R., Raab K., Schnitzler P., Janssen W., Darai G.;
 RT "Identification of the gene encoding the major capsid protein of
 RL insect iridescent virus type 6 by polymerase chain reaction.";
 RL J. Gen. Virol. 74:873-879(1993).
 [8]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94353641; PubMed=8073636;
 RA Sonntag K.C., Schnitzler P., Koonin E.V., Darai G.;
 RT "Chilo iridescent virus encodes a putative helicase belonging to a
 RL distinct family within the 'DEAD/H' superfamily: implications for the
 evolution of large DNA viruses.";
 RL Virus Genes 8:151-158(1994).
 [10]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95213160; PubMed=7998884;
 RA Sonntag K.C., Schnitzler P., Janssen W., Darai G.;
 RT "Identification of the primary structure and the coding capacity of
 RL the genome of insect iridescent virus type 6 between the genome
 coordinates 0.310 and 0.347 (7990 bp).";
 RL Intervirology 37:287-297(1994).
 [11]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94292906; PubMed=8021587;
 RA Schnitzler P., Sonntag K.C., Muller M., Janssen W., Bugert J.J.,
 RT "Insect iridescent virus type 6 encodes a polypeptide related to the
 RL largest subunit of eukaryotic RNA polymerase II.";
 RL J. Gen. Virol. 75:1557-1567(1994).

RN [12]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98141693; PubMed=9482589;
 RA Bahr U., Tidona C.A., Darai G.;
 RT "The DNA sequence of Chilo iridescent virus between the genome
 RL coordinates 0.101 and 0.391; similarities in coding strategy between
 RT insect and vertebrate iridoviruses.";
 RL Virus Genes 15:235-245(1997).
 [13]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99125223; PubMed=9926400;
 RA Muller K., Tidona C.A., Bahr U., Darai G.;
 RT "Identification of a thymidylate synthase gene within the genome of
 RL Chilo iridescent virus.";
 RL Virus Genes 17:243-258(1998).
 [14]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99383793; PubMed=10456793;
 RA Muller K., Tidona C.A., Darai G.;
 RT "Identification of a gene cluster within the genome of Chilo
 RL iridescent virus encoding enzymes involved in viral DNA replication
 and processing.";
 RL Virus Genes 18:243-264(1999).
 [15]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21342589; PubMed=11448171;
 RA Jakob N.J., Muller K., Bahr U., Darai G.;
 RT "Analysis of the First Complete DNA Sequence of an Invertebrate
 RL Iridovirus: Coding Strategy of the Genome of Chilo Irdescent Virus.";
 RL Virology 286:182-196(2001).
 [16]
 RP SEQUENCE FROM N.A.
 RA Jakob N.J., Mueller K., Bahr U., Darai G.;
 RT Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF303741; AAK82216.1; -.
 DR InterPro; IPR004274; NIF.
 DR Pfam; PF03031; NIF; 1.
 DR SMART; SM00577; CPDc; 1.
 SQ SEQUENCE 182 AA; 22021 MW; 5089911D0C4526E09 CRC64;
 Query Match 17.5%; Score 7; DB 12; Length 182;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 6 KKDPSRK 12
 Db 85 KKDPSRK 91
 RESULT 7
 Q8DT68 PRELIMINARY; PRT; 285 AA.
 AC Q8DT68;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Putative transcriptional regulator.
 GN RCG OR SMW.1509.
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_Taxid=1309;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UA159 / ATCC 700610 / Serotype C;
 RX MEDLINE=22295063; PubMed=12397186;
 RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
 RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
 RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
 RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
 RT pathogen.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
 DR EMBL; AE014983; AAN59160.1; -.

KW Complete proteome.
SQ SEQUENCE 285 AA; 33899 MW; D02B6AA2992AB6A5 CRC64;

Query Match
Best Local Similarity 17.5%; Score 7; DB 16; Length 285;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 TTSKMT 37
DB 43 TTSKMT 49

RESULT 8
Q22192 PRELIMINARY; PRT; 406 AA.
AC Q22192;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE T05A10.3 protein.
GN T05A10.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Sulston J.E.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL; 268108; CAA92135.1;-
DR WormPep; T05A10.3; CE03635.
DR InterPro; IPR001534; Worm_Family_2.
DR Pfam; PF01060; DUF290; 1.
SQ SEQUENCE 406 AA; 48925 MW; 8922CF21E33E9A01 CRC64;

Query Match
Best Local Similarity 17.5%; Score 7; DB 5; Length 406;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 KYEKDP 9
DB 369 KYEKDP 375

RESULT 9
Q9S233 PRELIMINARY; PRT; 406 AA.
AC Q9S233;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE AVR9 elicitor response like protein.
GN F10M23.280 OR AT4G26940.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucoside II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Lecharny A., Chefor F., Krivitzky M., Kreis M.,
RA Hohnsbeil J., Mewes H.W., Mayer K.F.X., Schueller C.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;

RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Lecharny A., Chefor F., Krivitzky M., Kreis M., Mewes H.W.,
RA Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL035440; CAB36540.1;-
DR EMBL; AL161566; CAB79549.1;-
DR InterPro; IPR002659; Glyco_trans_31.
DR Pfam; PF01762; Galactosyl_T; 1.
SQ SEQUENCE 406 AA; 45892 MW; 1141132B296F8B5 CRC64;

Query Match
Best Local Similarity 17.5%; Score 7; DB 10; Length 406;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 DKTISKL 35
DB 97 DKTISKL 103

RESULT 10
Q8LEJ9 PRELIMINARY; PRT; 407 AA.
AC Q8LEJ9;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE AVR9 elicitor response-like protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucoside II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Haas B.J., Volkovskiy N., Town C.D., Troukhan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT "Full-length messenger RNA sequences greatly improve genome
annotation."
RL Genome Biol. 0:0-0(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Broeyer V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RT "Full-length cDNA from Arabidopsis thaliana."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY085383; AAM62612.1;-
DR InterPro; IPR002659; Glyco_trans_31.
DR Pfam; PF01762; Galactosyl_T; 1.
SQ SEQUENCE 407 AA; 46020 MW; 756C0F866E1C952 CRC64;

Query Match
Best Local Similarity 17.5%; Score 7; DB 10; Length 407;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 DKTISKL 35
DB 98 DKTISKL 104

RESULT 11
Q81787 PRELIMINARY; PRT; 616 AA.
AC Q81787;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Erythrocyte binding antigen region II (fragment).

```

GN EBA-140.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5633;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIG203;
RA Baum J., Thomas A.W., Conway D.J.;
RT "Natural Selection on Erythrocyte Binding Proteins of Plasmodium
   falciparum.";
RL Submitted (MAR-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; AJ348840; CAD27565.1; -.
FT NON_TER
FT NON_TER
SQ SEQUENCE 616 AA; 73442 MW; C50CE1C56E05F54 CRC64;

Query Match 17.5%; Score 7; DB 5; Length 616;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 12 KKEWMEN 18
Db 139 KKEWMEN 145

RESULT 12
ID 001441 PRELIMINARY; PRT; 779 AA.
AC 001441.
DT 01-JUL-1997 (TRENBLREL. 04, Created)
DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)
DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)
DE Hypothetical 87.1 kDa protein.
GN B0261.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
   investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Scheet P., Gattung S.;
RT "The sequence of C. elegans comid B0261.";
RL Submitted (APR-1997) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT "Direct Submission";
RL Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 MYB-LIKE DOMAIN.
DR EMBL; U97016; AAB52350.2; -.
DR WormPep; B0261.1; CE295931.
DR InterPro; IPR001005; Myb_DNA_binding.
DR Pfam; PF00249; myb_DNA-binding; 1.
DR SMART; SM00717; SANT; 1.
KW Hypothetical protein; DNA-binding; Nuclear protein.
SQ SEQUENCE 779 AA; 87113 MW; 64B5E808E5A9674 CRC64;

Query Match 17.5%; Score 7; DB 5; Length 779;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 32 ISKLMTE 38

```

```

Db 330 ISKLMTE 336

RESULT 13
ID 0980C0 PRELIMINARY; PRT; 913 AA.
AC 0980C0.
DT 01-OCT-2001 (TRENBLREL. 18, Created)
DT 01-OCT-2001 (TRENBLREL. 18, Last sequence update)
DT 01-OCT-2001 (TRENBLREL. 18, Last annotation update)
DE ATP-dependent helicase.
GN SSO0394.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=11427726;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
   Awayez M.J., Chan-Weiner C.C.-Y., Clausen I.G., Curtis B.A.,
   De Moors A., Erasuo G., Fletcher C., Gordon P.M.K.,
   Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
   Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
   Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
   Garrett R.A., Ragan M.A., Senses C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
DR EMBL; AE006673; AAK40723.1; -.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
KW Helicase; Complete proteome.
SQ SEQUENCE 913 AA; 104861 MW; 99B596770227FC15 CRC64;

Query Match 17.5%; Score 7; DB 17; Length 913;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 PKVEKXD 8
Db 222 PKVEKXD 228

RESULT 14
ID 065580 PRELIMINARY; PRT; 1029 AA.
AC 065580.
DT 01-AUG-1998 (TRENBLREL. 07, Created)
DT 01-AUG-1998 (TRENBLREL. 07, Last sequence update)
DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)
DE Receptor protein kinase - like protein (Receptor protein kinase-like
   protein).
GN M3E9.30 OR A74G26540.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M. Jaller C., Portetelle D., Hohnseel J.,
   Beyer M., Vandenbol M., Jaller C., Portetelle D., Hohnseel J.,
   Mewes H.W., Mayer K., Schueller C.;
RL Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=M. Jaller C., Portetelle D., Mewes H.W., Lemcke K.,
   Mayer K.F.X.;

```

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Zimmermann W., Gruenisen A., Wambutt R., Kalicki J., Woldmann P.,
 RA Smith A., Mewes H.W., Lemcke K., Mayer K.F.X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP PARTIAL SEQUENCE FROM N.A.
 RA Volckaert G., Grymoprez B., Voet M., Robben J., Mewes H.W.,
 RA Lemcke K., Mayer K.F.X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; AL022223; CA18216.1; -;
 DR EMBL; AL161565; CAB79509.1; -;
 DR InterPro; IPR001611; LRR;
 DR InterPro; IPR007090; LRR_plant;
 DR InterPro; IPR000719; Prot_Kinase;
 DR InterPro; IPR002290; Ser_thr_kinase;
 DR Pfam; PF00560; LRR; 19;
 DR Pfam; PF00069; pkinase; 1;
 DR PRINTS; PR00019; LEURICHRPT.
 DR Prodom; PD000001; Prot_kinase; 1;
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1;
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1;
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1;
 KW ATP-binding; kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 1029 AA; 112970 MW; 0BE6FCF91D21386B CRC64;

Query Match 17.5%; Score 7; DB 10; Length 1029;
 Best Local Similarity 100.0%; Pred. No. 57;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 EKKDPSR 11
 Db 925 EKKDPSR 931

RESULT 15
 O9SIP2 PRELIMINARY; PRT; 1037 AA.
 AC O9SIP2; 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Putative importin (Nuclear transport factor) protein.
 GN ATG31660.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=20083487; PubMed=10617197;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
 RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
 RA Cronin L.A., Shen W., Vanaken S.E., Unayam L., Tallon L.J., Gill J.E.,
 RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
 RA Copenhaver G.P., Preuss D., Niernman W.C., White O., Eisen J.A.,
 RA Salzberg S.L., Fraser C.M., Venter J.C.;
 RA "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 thaliana";
 RT Nature 402:761-768(1999).
 RL Nature 402:761-768(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;

RA Lin X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC007071; AAD24843.1; -;
 DR InterPro; IPR001494; Importinb_N.
 DR Pfam; PF03810; IEN_NT; 1;
 DR PROSITE; PS0166; IMPORTIN_B_NT; 1;
 SQ SEQUENCE 1037 AA; 118966 MW; 4DE90DCCF7B3C6DA CRC64;

Query Match 17.5%; Score 7; DB 10; Length 1037;
 Best Local Similarity 100.0%; Pred. No. 57;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 31 TTSKMT 37
 Db 721 TTSKMT 727

Search completed: July 24, 2003, 12:25:52
 Job time : 15.2987 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 11:44:42 ; Search time 13.0015 Seconds

(without alignments)
488.334 Million cell updates/sec

Title: US-09-991-681-30

Perfect score: 215
Sequence: 1 SPKVEKKDPSRKKEMWENAKIYTMADKTIKLTETK 40

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	215	100.0	40	23	AAO19168 Human prostate-spe
2	215	100.0	40	23	AAO19168 Human prostate-spe
3	215	100.0	518	19	AAW85472 Human prostate-enco
4	215	100.0	518	23	AAO19165 Human prostate-spe
5	215	100.0	518	23	AAW85472 Human prostate-spe
6	215	100.0	1807	22	ABG09728 Novel human diagno
7	215	100.0	1839	23	ABP64835 Human protein SEQ
8	215	100.0	1982	22	ABG09731 Novel human diagno
9	59.5	27.7	388	22	AAU36030 Helicobacter pylori

10	57	26.5	664	22	AAO6034 Mouse oocyte prote
11	57	26.5	682	24	AAE32624 Mouse peptidylargi
12	56.5	26.3	388	19	AAW8375 H. pylori GPO 113
13	56.5	26.3	388	22	AAU35865 Helicobacter pylori
14	56	26.0	219	21	AAO44032 Arabidopsis thalia
15	56	26.0	219	21	AAO44032 Arabidopsis thalia
16	56	26.0	221	21	AAO44031 Arabidopsis thalia
17	56	26.0	221	21	AAO44031 Arabidopsis thalia
18	56	26.0	227	21	AAO44036 Arabidopsis thalia
19	56	26.0	241	21	AAO44037 Arabidopsis thalia
20	56	26.0	304	23	ABE91690 Arabidopsis thalia
21	56	26.0	317	21	AAO50357 Arabidopsis thalia
22	56	26.0	326	21	AAO50356 Arabidopsis thalia
23	56	26.0	328	21	AAO50355 Arabidopsis thalia
24	56	26.0	542	23	ABP73803 Arabidopsis thalia
25	55	25.6	351	20	AAW82796 Candida albicans e
26	55	25.6	351	20	ABP71036 Tobacco TobRD2 pro
27	55	25.6	654	23	ABE93787 Tobacco quiniolate
28	53	24.7	141	22	ABG04874 Herbicidally activ
29	53	24.7	270	21	AAO18382 Novel human diagno
30	53	24.7	338	21	AAO18381 Arabidopsis thalia
31	53	24.7	373	23	AAU11287 Human transducin p
32	53	24.7	410	23	ABE93088 Herbicidally activ
33	53	24.7	423	23	ABE93765 Herbicidally activ
34	53	24.7	462	22	AAO81158 Mycobacterium tube
35	53	24.7	488	22	ABE60032 Drosophila melanog
36	53	24.7	528	22	AAE93594 Human protein sequ
37	53	24.7	656	23	AAE16259 Human kinase PKIN-
38	53	24.7	900	23	ABE93716 Herbicidally activ
39	53	24.7	909	22	AAU03554 Human protein kina
40	53	24.7	911	24	ABU11054 Human protein NOVI
41	53	24.7	915	24	ABU11053 Human protein NOVI
42	53	24.7	2527	24	ABU08105 Human kinase and p
43	52.5	24.4	232	21	AAE83330 Human secreted pro
44	52.5	24.4	234	22	AAE97771 Hevea brasiliensis
45	52.5	24.4	234	22	AAE30549 An isopenyl dipho

ALIGNMENTS

RESULT 1
AAO19168 standard; Protein: 40 AA.
XX AAO19168;
XX AC
XX 27-NOV-2002 (first entry)
DT
XX Human prostate-specific PS18 protein fragment #4.
DE
XX Human, prostate; prostate-specific sequence; prostate cancer; PS18;
KW EST; expressed sequence tag; cytosolic; gene therapy.
XX
XX Homo sapiens.
OS
XX US2002086316-A1.
XX PN
XX 04-JUL-2002.
XX
XX 26-NOV-2001; 2001US-0991681.
XX PF
XX 23-APR-1998; 98US-0065383.
XX PR 23-APR-1997; 97US-0842385.
XX
XX (BILU/) BILINGEL P A.
XX (COHE/) COHEN M.
XX (COLP/) COLPITS T L.
XX (FRIE/) FRIEDMAN P N.
XX (GORD/) GORDAN J.
XX (GRAN/) GRANADOS E N.
XX (HODG/) HODGES S C.
XX (KLAS/) KLAS M R.

PA (KRAT/) KRATOCHVIL J D.
 PA (ROBE/) ROBERTS-RAPP L.
 PA (RUSS/) RUSSELL J C.
 PA (STRO/) STROUPE S D.
 XX
 PI Billingsel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;
 PI Granados EN, Hodges SC, Klasse MR, Kratochvil JD, Roberts-Rapp L;
 PI Russell JC, Stroupe SD;
 XX
 DR WPI; 2002-665429/71.
 XX
 PT Novel PS118 polypeptide for detecting, diagnosing, staging, monitoring,
 PT prognosticating, preventing, treating, or determining predisposition of
 PT individual to diseases and conditions of prostate, e.g. prostate
 PT cancer -
 XX
 PS Claim 17; Page 44; 589p; English.
 XX
 CC The present invention relates to a number of prostate-specific sequences
 CC derived from the human PS118 gene. These can be used in the detection,
 CC monitoring and treatment of prostate diseases, particularly prostate
 CC cancer. The present sequence is a PS118 protein fragment of the
 CC invention. The coding sequences of the invention were isolated from a
 CC prostate tissue expressed sequence tag (EST) library.
 XX
 SQ Sequence 40 AA;
 Query Match 100.0%; Score 215; DB 23; Length 40;
 Best Local Similarity 100.0%; Pred. No. 2.8e-21;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SPKVEKKDPSRKKEWENAGNKITYTMAADKITSKLMTEYK 40
 DB 1 SPKVEKKDPSRKKEWENAGNKITYTMAADKITSKLMTEYK 40
 XX
 RESULT 2
 ID AAM50812
 AC AAM50812 standard; Protein; 40 AA.
 XX
 AC AAM50812;
 XX
 DT 01-MAY-2002 (first entry)
 XX
 DE PS118 prostate marker immunogenic polypeptide.
 XX
 KW PS118; prostate; marker; prostate cancer; tumour; metastasis;
 KW benign prostatic hyperplasia; prostatic intraepithelial neoplasia;
 KW prostatic; human; diagnosis; therapy; vaccine; immunogen.
 XX
 OS Homo sapiens.
 XX
 PN US2001055758-A1.
 XX
 PD 27-DEC-2001.
 XX
 PF 23-APR-1998; 98US-0065383.
 XX
 PR 23-APR-1997; 97US-0842385.
 XX
 PA (BILL/) BILLING-MEDEL P A.
 PA (COHE/) COHEN M.
 PA (COPL/) COPLITTIS T L.
 PA (FRIE/) FRIEDMAN P N.
 PA (GORD/) GORDON J.
 PA (GRAN/) GRANADOS E N.
 PA (HODG/) HODGES S C.
 PA (KLAS/) KLASSE M R.
 PA (KRAT/) KRATOCHVIL J D.
 PA (ROBE/) ROBERTS-RAPP L.
 XX
 PI Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;
 PI Granados EN, Hodges SC, Klasse MR, Kratochvil JD, Roberts-Rapp L;

PI Russell JC, Stroupe SD;
 XX
 DR WPI; 2002-187683/24.
 XX
 PT Detecting presence of target PS118 polynucleotide in test sample,
 PT useful for detecting, diagnosing, staging, monitoring, prognosticating,
 PT preventing or treating or determining predisposition to prostate
 PT disease -
 XX
 PS Claim 17; Page 44; 579p; English.
 XX
 CC The present sequence is that of an immunogenic polypeptide
 CC comprising amino acids 332-371 of human prostate-specific PS118
 CC polypeptide (see AAM50809). A PS118 consensus sequence (see
 CC ABA91651) is found at least 12 times more often in prostate than
 CC in non-prostate tissue. PS118 polypeptides, polynucleotides,
 CC antibodies, agonists and inhibitors are useful for detecting,
 CC diagnosing, staging, monitoring, prognosticating, preventing and
 CC treating (including by genetic immunisation), or determining the
 CC predisposition of an individual to, diseases and conditions of the
 CC prostate, such as benign prostatic hyperplasia, prostaticitis,
 CC prostatic intraepithelial neoplasia, prostate cancer, tumours and
 CC metastases. The PS118 polypeptides can be produced by expression
 CC of PS118 polynucleotides in transfected host cells, and
 CC immunogenic peptides are useful for raising PS118-specific
 CC antibodies of diagnostic use. The methods and reagents of the
 CC invention may provide an early means of detecting diseases of the
 CC prostate and may also provide new markers which can differentiate
 CC between the clinically important and unimportant prostate cancers
 CC without the use of surgery.
 XX
 SQ Sequence 40 AA;
 Query Match 100.0%; Score 215; DB 23; Length 40;
 Best Local Similarity 100.0%; Pred. No. 2.8e-21;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SPKVEKKDPSRKKEWENAGNKITYTMAADKITSKLMTEYK 40
 DB 1 SPKVEKKDPSRKKEWENAGNKITYTMAADKITSKLMTEYK 40
 XX
 RESULT 3
 ID AAW85472
 AC AAW85472 standard; Protein; 518 AA.
 XX
 AC AAW85472;
 XX
 DT 25-MAR-2003 (updated)
 DT 25-FEB-1999 (first entry)
 XX
 DE PS118 protein encoded by consensus sequence.
 XX
 KW EST clone; PS118; prostate tumour tissue; prostatic disease; cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO9848054-A1.
 XX
 PD 29-OCT-1998.
 XX
 PF 23-APR-1998; 98WO-US08239.
 XX
 PR 23-APR-1997; 97US-0842385.
 XX
 PA (ABBO) ABBOTT LAB.
 XX
 PI Billing-medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;
 PI Granados EN, Hodges SC, Klasse MR, Kratochvil JD, Robertstapp L;
 PI Russell JC, Stroupe SD;
 XX
 DR WPI; 1998-610000/51.
 DR N-PSDB; AAW82812.

XX New P118 nucleic acid and proteins - used for diagnosis and
PT treatment of prostatic disease, especially cancer, and also for drug
PT screening
XX
PS Claim 17; Page 93-94; 117pp; English.
XX
PS The present sequence is encoded by consensus P5118 sequence derived from
CC overlapping EST (expressed sequence tag) clones AAV82803-10. The clones
CC were identified from a cDNA library made from prostate tumour tissue.
CC Recombinant P5118 protein is used to detect P5118-specific antibodies,
CC to raise antibodies for detection of P5118 antigens, to screen for
CC specific binding agents (potential therapeutics, and to isolate specific
CC antibodies from serum. Detection of P5118 protein or nucleic acid, which
CC are prostate related, and altered or elevated in prostatic disease, is
CC used for detection, diagnosis, staging, monitoring and prognosis of
CC prostatic disease, particularly cancer, and to identify subjects at
CC risk.
CC (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ Sequence 518 AA;
XX
Query Match 100.0%; Score 215; DB 19; Length 518;
Best Local Similarity 100.0%; Pred. No. 4.9e-20;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 SPKVEKKDPSRKKEWENAGNKIYTMADKTISKLMTEYK 40
DB 332 SPKVEKKDPSRKKEWENAGNKIYTMADKTISKLMTEYK 371
XX
RESULT 4
ID AAO19165 standard; Protein; 518 AA.
XX
AC AAO19165;
XX
DT 27-NOV-2002 (first entry)
XX
DE Human prostate-specific P5118 protein fragment #1.
XX
KW Human; prostate; prostate-specific sequence; prostate cancer; P5118;
KW EST; expressed sequence tag; cytostatic; gene therapy.
XX
OS Homo sapiens.
XX
PN US2002086316-A1.
XX
PD 04-JUL-2002.
XX
PF 26-NOV-2001; 2001US-0991681.
XX
PR 23-APR-1998; 98US-0065383.
PR 23-APR-1997; 97US-0842385.
XX
PA (BILL/) BILLINGEL P A.
PA (COHE/) COHEN M.
PA (COLP/) COLPITTS T L.
PA (FRIE/) FRIEDMAN P N.
PA (GORD/) GORDAN J.
PA (GRAN/) GRANADOS E N.
PA (HODG/) HODGES S C.
PA (KLAS/) KLAAS M R.
PA (KRAT/) KRATOCHVIL J D.
PA (ROBE/) ROBERTS-RAPP L.
PA (RUSSE/) RUSSELL J C.
PA (STRO/) STROUPE S D.
XX
PI Billingsel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;
PI Granados EN, Hodges SC, Klaas MR, Kratochvil JD, Roberts-Rapp L;
PI Russell JC, Stroupe SD;
XX WPI; 2002-665429/71.
DR

XX Novel P5118 polypeptide for detecting, diagnosing, staging, monitoring,
PT prognosticating, preventing, treating, or determining predisposition of
PT individual to diseases and conditions of prostate, e.g. prostate
PT cancer -
XX
PS Claim 17; Page 42-43; 58pp; English.
XX
PS The present invention relates to a number of prostate-specific sequences
CC derived from the human P5118 gene. These can be used in the detection,
CC monitoring and treatment of prostate diseases, particularly prostate
CC cancer. The present sequence is a P5118 protein fragment of the
CC invention. The coding sequences of the invention were isolated from a
XX prostate tissue expressed sequence tag (EST) library.
XX
SQ Sequence 518 AA;
XX
Query Match 100.0%; Score 215; DB 23; Length 518;
Best Local Similarity 100.0%; Pred. No. 4.9e-20;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 SPKVEKKDPSRKKEWENAGNKIYTMADKTISKLMTEYK 40
DB 332 SPKVEKKDPSRKKEWENAGNKIYTMADKTISKLMTEYK 371
XX
RESULT 5
ID AAM50809 standard; Protein; 518 AA.
XX
AC AAM50809;
XX
DT 01-MAY-2002 (first entry)
XX
DE P5118 prostate marker partial sequence.
XX
KW P5118; prostate; marker; prostate cancer; tumour; metastasis;
KW benign prostatic hyperplasia; prostatic intraepithelial neoplasia;
KW prostatitis; human; diagnosis; therapy; vaccine.
XX
OS Homo sapiens.
XX
PN US2001055758-A1.
XX
PD 27-DEC-2001.
XX
PF 23-APR-1998; 98US-0065383.
XX
PR 23-APR-1997; 97US-0842385.
XX
PA (BILL/) BILLING-MEDEL P A.
PA (COHE/) COHEN M.
PA (COLP/) COLPITTS T L.
PA (FRIE/) FRIEDMAN P N.
PA (GORD/) GORDON J.
PA (GRAN/) GRANADOS E N.
PA (HODG/) HODGES S C.
PA (KLAS/) KLAAS M R.
PA (KRAT/) KRATOCHVIL J D.
PA (ROBE/) ROBERTS-RAPP L.
XX
PI Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;
PI Granados EN, Hodges SC, Klaas MR, Kratochvil JD, Roberts-Rapp L;
PI Russell JC, Stroupe SD;
XX WPI; 2002-187683/24.
XX N-PSDB; ABA91651.
DR
XX
PT Detecting presence of target P5118 polynucleotide in test sample,
PT useful for detecting, diagnosing, staging, monitoring, prognosticating,
PT preventing or treating or determining predisposition to prostate
PT disease -
XX

PS Claim 17; Page 42-43; 57pp; English.

XX The present sequence is that of a human prostate-specific PS118
CC polypeptide, as predicted from a partial consensus cDNA sequence
CC (see ABA91651), and lacking the N-terminal region. The PS118
CC consensus sequence is found at least 12 times more often in
CC prostate than in non-prostate tissue. PS118 polypeptides,
CC including derivatives of the present sequence, polynucleotides,
CC antibodies, agonists and inhibitors are useful for detecting,
CC diagnosing, staging, monitoring, prognosticating, preventing and
CC treating (including by genetic immunisation), or determining the
CC predisposition of an individual to, diseases and conditions of the
CC prostate, such as benign prostatic hyperplasia, prostatitis,
CC prostatic intraepithelial neoplasia, prostate cancer, tumours and
CC metastases. The PS118 polypeptides can be produced by expression
CC of PS118 polynucleotides in transfected host cells. The methods
CC and reagents of the invention may provide an early means of
CC detecting diseases of the prostate and may also provide new markers
CC which can differentiate between the clinically important and
CC unimportant prostate cancers without the use of surgery.

XX Sequence 518 AA;

SQ

Query Match 100.0%; Score 215; DB 23; Length 518;
Best Local Similarity 100.0%; Pred. No. 4,9e-20;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SPKVEKKDPSRKKEWENAGNKIYTMADKTIISKLMTEYK 40
DB 332 SPKVEKKDPSRKKEWENAGNKIYTMADKTIISKLMTEYK 371

RESULT 6
ABG09728
ID ABG09728 standard; Protein; 1807 AA.

XX ABG09728;
XX
XX 13-FEB-2002 (first entry)
XX
XX Novel human diagnostic protein #9719.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
XX
XX 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YF;
XX
XX WPI; 2001-639362/73.
XX
XX N-PSDB; AAS73915.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity -
XX
XX Claim 20; SEQ ID No 40087; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1807 AA;

SQ

Query Match 100.0%; Score 215; DB 22; Length 1807;
Best Local Similarity 100.0%; Pred. No. 2e-19;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SPKVEKKDPSRKKEWENAGNKIYTMADKTIISKLMTEYK 40
DB 1621 SPKVEKKDPSRKKEWENAGNKIYTMADKTIISKLMTEYK 1660

RESULT 7
ABP64835
ID ABP64835 standard; Protein; 1839 AA.

XX ABP64835;
XX
XX 25-FEB-2003 (first entry)
XX
XX Human protein SEQ ID 495.
XX
XX
XX Human; expressed sequence tag; EST;
XX haematopoietic disorder; central nervous system disease; viral infection;
XX peripheral nervous system disease; non-healing wound; infectious disease;
XX immune deficiency; immune disorder; bacterial infection; allergy; cancer;
XX fungal infection; autoimmune disorder; coagulation disorder; noctropic;
XX anti-allergic; anti-inflammatory; immunosuppressive; neuroprotective;
XX cytostatic; haemostatic; virucide; antibacterial; fungicide;
XX immunostimulant; cerebroprotective.

XX
XX Homo sapiens.
XX
XX WO200259260-A2.
XX
XX 01-AUG-2002.
XX
XX 16-NOV-2001; 2001WO-US42950.
XX
XX 17-NOV-2000; 2000US-0714936.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YF, Goodrich RW, Liu C, Zhou P, Aouni V, Zhang J, Zhao QA;
XX Ren F, Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX
XX WPI; 2002-590824/63.
XX
XX N-PSDB; ABQ99421.
XX
XX New isolated polynucleotide, useful in research, diagnostic or
XX therapeutic methods, e.g. preventing or treating disorders involving
XX aberrant protein expression or biological activity -
XX
XX Claim 20; SEQ ID 495; 394pp; English.

CC The present invention relates to novel human coding sequences
CC (ABG09268-ABG09608) and proteins (ABP64682-ABP65022). The sequences are
CC useful in therapeutic, diagnostic and research methods. The
CC polynucleotides may be used in the field of molecular biology as
CC hybridisation probes, primers for PCR, for chromosome and gene mapping,
CC for the recombinant production of protein, or in generation of anti-sense
CC DNA or RNA. The polynucleotides are useful in diagnostics as expressed
CC sequence tags (ESTs) for identifying expressed genes or for physical
CC mapping of the human genome. The proteins may be used as molecular weight
CC markers, or as nutritional sources or supplements. The proteins may be
CC used to maintain and expand cell population in a totipotential or
CC pluripotential state useful for re-engineering damaged or diseased
CC tissues, transplantation, manufacture of bio-pharmaceuticals or the
CC development of bio-sensors. The polynucleotides and proteins are useful
CC for preventing, treating or ameliorating disorders involving aberrant
CC protein expression or biological activity, e.g. haematopoietic disorders,
CC central/peripheral nervous system diseases, mechanical and traumatic
CC disorders, non-healing wounds, immune deficiencies and disorders,
CC infectious diseases caused by viral, bacterial or fungal infection,
CC autoimmune disorders, allergic reactions and conditions, coagulation
CC disorders, or cancer. The polynucleotide sequences of the invention were
CC assembled from ESTs isolated mainly by sequencing by hybridisation, and
CC in some cases, sequences obtained from one or more public databases.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 1839 AA;

Query Match 100.0%; Score 215; DB 23; Length 1839;
Best Local Similarity 100.0%; Pred. No. 2.1e-19;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPKVEKKDPSRKKEWENAGNKIYTMADKTIISKLMTEYK 40
Db 1653 SPKVEKKDPSRKKEWENAGNKIYTMADKTIISKLMTEYK 1692

RESULT 8
ABG09731
ID ABG09731 standard; Protein; 1982 AA.

AC ABG09731;
XX
DT 13-FEB-2002 (first entry)
XX

DE Novel human diagnostic protein #9722.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

OS WO200175067-A2.

PN 11-OCT-2001.

PD 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

PA Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

DR N-PSDB; AAS73918.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity

XX Claim 20; SEQ ID No 40090; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. ABG00010-ABG30377 represent novel human
XX diagnostic amino acid sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 1982 AA;

Query Match 100.0%; Score 215; DB 22; Length 1982;
Best Local Similarity 100.0%; Pred. No. 2.2e-19;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPKVEKKDPSRKKEWENAGNKIYTMADKTIISKLMTEYK 40
Db 1735 SPKVEKKDPSRKKEWENAGNKIYTMADKTIISKLMTEYK 1774

RESULT 9
AAU36030
ID AAU36030 standard; Protein; 388 AA.

AC AAU36030;

XX 14-FEB-2002 (first entry)

DE Helicobacter pylori cellular proliferation protein #243.

KW Antisense; prokaryotic cellular proliferation protein;
KW antibiotic; antibacterial; drug design.

XX Helicobacter pylori.

OS WO200170955-A2.

PN 27-SEP-2001.

PD 21-MAR-2001; 2001WO-US09180.

PR 21-MAR-2000; 2000US-191078P.

PR 23-MAY-2000; 2000US-206848P.

PR 26-MAY-2000; 2000US-207727P.

PR 23-OCT-2000; 2000US-242578P.

PR 27-NOV-2000; 2000US-253625P.

PR 22-DEC-2000; 2000US-257931P.

PR 16-FEB-2001; 2001US-269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zykkind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX WPI; 2001-611495/70.

DR N-PSDB; AAS53889.

PT New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -
 XX
 PS Example 3; Seq ID No 11623; 511pp; English.
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
 CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 388 AA;
 Query Match 27.7%; Score 59.5; DB 22; Length 388;
 Best Local Similarity 37.1%; Pred. No. 15;
 Matches 13; Conservative 4; Mismatches 13; Indels 5; Gaps 1;
 QY 11 RKKEWENAGNKITYTMADK-----TISKLTMEYK 40
 DB 217 RUKLWMSNLQNSLFTLLPDKLNNALRISDLPSYQ 251
 RESULT 10
 AAE06034
 ID AAE06034 standard; Protein; 664 AA.
 AC AAE06034;
 XX
 XX
 DT 25-SEP-2001 (first entry)
 DE Mouse oocyte protein 5 (MOP5).
 XX
 KM Mouse oocyte protein 5; MOP5; egg specific surface protein; immunogen;
 KM vaccine; fertility; contraceptive; active immunisation; sterilisation.
 XX
 OS Mus musculus.
 XX
 PN WO200153339-A2.
 PD 26-JUL-2001.
 XX
 PF 19-JAN-2001; 2001WO-US01718.
 XX
 PR 20-JAN-2000; 2000US-0177123.
 XX
 PA (UUYI-) UNIV VIRGINIA PATENT FOUND.
 XX
 PI Herr JC, Coomrod SA, Wright P;
 XX
 DR WPI; 2001-465367/50.
 DR N-PSDB; AAD11617, AAD11618.
 XX
 PT New egg-specific surface proteins, useful as immunogens in a vaccine
 PT preparation for modulating fertility, particularly useful for producing
 PT antibodies for temporary, reversible contraception methods -
 XX
 PS Claim 1; Page 32-35; 41pp; English.
 CC The present sequence is mouse oocyte protein 5 (MOP5). The

CC Present invention relates to egg specific surface proteins
 CC (antigens) and nucleic acids encoding them. The proteins of the
 CC invention are involved in egg-sperm binding and fusion. They are
 CC useful as immunogens in vaccine preparation for modulating fertility.
 CC In particular, the proteins are useful for producing antibodies
 CC which are useful for temporary, reversible contraception methods.
 CC The contraceptive vaccine is especially useful for controlling
 CC fertility in human beings or agriculturally important livestock,
 CC e.g. cattle or pig. The invention also relates to the use of
 CC antibodies against such antigens for active immunisation or
 CC sterilisation of female animals.
 CC
 SQ Sequence 664 AA;
 Query Match 26.5%; Score 57; DB 22; Length 664;
 Best Local Similarity 33.3%; Pred. No. 59;
 Matches 10; Conservative 9; Mismatches 11; Indels 0; Gaps 0;
 QY 7 KDPKRKKEWENAGNKITYTMADXTISKLM 36
 DB 318 EDPNRQSKWLQDEMAFCTQAPKTVSLIT 347
 RESULT 11
 AAE32624
 ID AAE32624 standard; Protein; 682 AA.
 AC AAE32624;
 XX
 XX
 DT 24-MAR-2003 (first entry)
 DE Mouse peptidylarginine deiminase 6 (PAD6).
 XX
 KM Peptidylarginine deiminase 6; PAD6; contraception; fertility disorder;
 KM gene therapy; mouse; enzyme.
 XX
 OS Mus musculus.
 XX
 PN WO200290531-A2.
 PD 14-NOV-2002.
 XX
 PF 25-APR-2002; 2002WO-EP04552.
 XX
 PR 03-MAY-2001; 2001EP-0201601.
 XX
 PA (ALKU) AKZO NOBEL NV.
 XX
 PI Gossen JA, Van Den Boogaart P;
 XX
 DR WPI; 2003-103517/09.
 DR N-PSDB; AAD50412.
 XX
 PT New polynucleotide comprising a nucleotide sequence that encodes
 PT peptidylarginine deiminase 6, useful as targets for male and female
 PT contraception and for screening small molecular weight modulators of
 PT the enzyme activity -
 XX
 PS Claim 3; Page 25-28; 35pp; English.
 XX
 CC The invention relates to peptidylarginine deiminase 6 (PAD6) polypeptides
 CC and polynucleotides. Polynucleotides of the invention are useful as
 CC targets for male and female contraception. Polypeptides of the invention
 CC are useful for screening small molecular weight modulators useful for
 CC controlling fertility disorders. They are also useful for screening
 CC therapeutic agents for fertility control related to the activity of PAD6.
 CC The invention is useful in gene therapy. The present sequence is mouse
 CC PAD6 protein.
 CC
 SQ Sequence 682 AA;
 Query Match 26.5%; Score 57; DB 24; Length 682;
 Best Local Similarity 33.3%; Pred. No. 61;

DT 14-FEB-2002 (first entry)

DT 18-OCT-2000 (First entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 55105.
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
PF 25-FEB-1999; 99US-0121825.
XX PR 05-MAR-1999; 99US-0123180.
PR PR 09-MAR-1999; 99US-0123548.
PR PR 23-MAR-1999; 99US-0125788.
PR PR 25-MAR-1999; 99US-0126264.
PR PR 29-MAR-1999; 99US-0126785.
PR PR 01-APR-1999; 99US-0127462.
PR PR 06-APR-1999; 99US-0128234.
PR PR 08-APR-1999; 99US-0128714.
PR PR 16-APR-1999; 99US-0129845.
PR PR 19-APR-1999; 99US-0130077.
PR PR 21-APR-1999; 99US-0130449.
PR PR 23-APR-1999; 99US-0130510.
PR PR 23-APR-1999; 99US-0130891.
PR PR 28-APR-1999; 99US-0131449.
PR PR 30-APR-1999; 99US-0132048.
PR PR 30-APR-1999; 99US-0132407.
PR PR 04-MAY-1999; 99US-0132484.
PR PR 05-MAY-1999; 99US-0132485.
PR PR 06-MAY-1999; 99US-0132486.
PR PR 07-MAY-1999; 99US-0132863.
PR PR 11-MAY-1999; 99US-0134256.
PR PR 14-MAY-1999; 99US-0134218.
PR PR 14-MAY-1999; 99US-0134219.
PR PR 14-MAY-1999; 99US-0134221.
PR PR 14-MAY-1999; 99US-0134370.
PR PR 18-MAY-1999; 99US-0134768.
PR PR 19-MAY-1999; 99US-0134941.
PR PR 20-MAY-1999; 99US-0135124.
PR PR 21-MAY-1999; 99US-0135353.
PR PR 24-MAY-1999; 99US-0135629.
PR PR 25-MAY-1999; 99US-0136021.
PR PR 27-MAY-1999; 99US-0136392.
PR PR 28-MAY-1999; 99US-0136782.
PR PR 01-JUN-1999; 99US-0137222.
PR PR 03-JUN-1999; 99US-0137528.
PR PR 04-JUN-1999; 99US-0137502.
PR PR 07-JUN-1999; 99US-0137724.
PR PR 08-JUN-1999; 99US-0138094.
PR PR 10-JUN-1999; 99US-0138540.
PR PR 10-JUN-1999; 99US-0138647.
PR PR 14-JUN-1999; 99US-0139119.
PR PR 16-JUN-1999; 99US-0139452.
PR PR 16-JUN-1999; 99US-0139453.
PR PR 17-JUN-1999; 99US-0139492.
PR PR 18-JUN-1999; 99US-0139454.
PR PR 18-JUN-1999; 99US-0139455.
PR PR 18-JUN-1999; 99US-0139456.
PR PR 18-JUN-1999; 99US-0139457.
PR PR 18-JUN-1999; 99US-0139458.
PR PR 18-JUN-1999; 99US-0139459.
PR PR 18-JUN-1999; 99US-0139460.
PR PR 18-JUN-1999; 99US-0139461.
PR PR 18-JUN-1999; 99US-0139462.
PR PR 18-JUN-1999; 99US-0139463.
PR PR 18-JUN-1999; 99US-0139750.
PR PR 18-JUN-1999; 99US-0139763.
PR PR 21-JUN-1999; 99US-0139817.

PR PR 22-JUN-1999; 99US-0139899.
PR PR 23-JUN-1999; 99US-0140353.
PR PR 23-JUN-1999; 99US-0140354.
PR PR 24-JUN-1999; 99US-0140695.
PR PR 28-JUN-1999; 99US-0140821.
PR PR 29-JUN-1999; 99US-0140991.
PR PR 30-JUN-1999; 99US-0141287.
PR PR 01-JUL-1999; 99US-0141842.
PR PR 01-JUL-1999; 99US-0142154.
PR PR 02-JUL-1999; 99US-0142055.
PR PR 06-JUL-1999; 99US-0142390.
PR PR 08-JUL-1999; 99US-0142803.
PR PR 09-JUL-1999; 99US-0142920.
PR PR 12-JUL-1999; 99US-0143547.
PR PR 13-JUL-1999; 99US-0143542.
PR PR 14-JUL-1999; 99US-0143624.
PR PR 15-JUL-1999; 99US-0144005.
PR PR 16-JUL-1999; 99US-0144085.
PR PR 16-JUL-1999; 99US-0144086.
PR PR 19-JUL-1999; 99US-0144325.
PR PR 19-JUL-1999; 99US-0144331.
PR PR 19-JUL-1999; 99US-0144332.
PR PR 19-JUL-1999; 99US-0144333.
PR PR 19-JUL-1999; 99US-0144334.
PR PR 19-JUL-1999; 99US-0144335.
PR PR 20-JUL-1999; 99US-0144352.
PR PR 20-JUL-1999; 99US-0144632.
PR PR 20-JUL-1999; 99US-0144884.
PR PR 21-JUL-1999; 99US-0144814.
PR PR 21-JUL-1999; 99US-0145086.
PR PR 21-JUL-1999; 99US-0145088.
PR PR 22-JUL-1999; 99US-0145085.
PR PR 22-JUL-1999; 99US-0145087.
PR PR 22-JUL-1999; 99US-0145089.
PR PR 22-JUL-1999; 99US-0145192.
PR PR 23-JUL-1999; 99US-0145145.
PR PR 23-JUL-1999; 99US-0145218.
PR PR 23-JUL-1999; 99US-0145224.
PR PR 26-JUL-1999; 99US-0145276.
PR PR 27-JUL-1999; 99US-0145913.
PR PR 27-JUL-1999; 99US-0145918.
PR PR 27-JUL-1999; 99US-0145919.
PR PR 28-JUL-1999; 99US-0145951.
PR PR 02-AUG-1999; 99US-0146386.
PR PR 02-AUG-1999; 99US-0146388.
PR PR 02-AUG-1999; 99US-0146389.
PR PR 03-AUG-1999; 99US-0147038.
PR PR 04-AUG-1999; 99US-0147204.
PR PR 04-AUG-1999; 99US-0147302.
PR PR 05-AUG-1999; 99US-0147192.
PR PR 05-AUG-1999; 99US-0147260.
PR PR 06-AUG-1999; 99US-0147303.
PR PR 06-AUG-1999; 99US-0147305.
PR PR 09-AUG-1999; 99US-0147493.
PR PR 09-AUG-1999; 99US-0147935.
PR PR 10-AUG-1999; 99US-0148171.
PR PR 11-AUG-1999; 99US-0148319.
PR PR 12-AUG-1999; 99US-0148341.
PR PR 13-AUG-1999; 99US-0148565.
PR PR 13-AUG-1999; 99US-0148684.
PR PR 16-AUG-1999; 99US-0149368.
PR PR 17-AUG-1999; 99US-0149175.
PR PR 18-AUG-1999; 99US-0149426.
PR PR 20-AUG-1999; 99US-0149722.
PR PR 20-AUG-1999; 99US-0149723.
PR PR 20-AUG-1999; 99US-0149929.
PR PR 23-AUG-1999; 99US-0149902.
PR PR 23-AUG-1999; 99US-0149930.
PR PR 25-AUG-1999; 99US-0150566.
PR PR 26-AUG-1999; 99US-0150884.
PR PR 27-AUG-1999; 99US-0151065.
PR PR 27-AUG-1999; 99US-0151066.
PR PR 27-AUG-1999; 99US-0151080.

PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 26.0%; Score 56; DB 21; Length 219;
Best Local Similarity 33.3%; Pred. No. 23;
Matches 10; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

Qy 9 PSRKKEWENANKITYMAADKTKSKMTE 38
Db 105 PSLQKQWYSVAGNHDRGNVEAQLSKVLTO 134

RESULT 15

AGS0369
ID AGS0369 standard; Protein; 219 AA.

XX AGS0369;

XX 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 63822.

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.

OS Arabidopsis thaliana.

XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123160.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130449.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 04-MAY-1999; 99US-0132407.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 06-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137750.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
XX 18-JUN-1999; 99US-0139462.
XX 18-JUN-1999; 99US-0139463.
XX 18-JUN-1999; 99US-0139750.
XX 18-JUN-1999; 99US-0139763.
XX 21-JUN-1999; 99US-0139817.
XX 22-JUN-1999; 99US-0139899.
XX 23-JUN-1999; 99US-0140353.
XX 23-JUN-1999; 99US-0140354.
XX 24-JUN-1999; 99US-0140685.
XX 28-JUN-1999; 99US-0140823.
XX 29-JUN-1999; 99US-0140991.
XX 30-JUN-1999; 99US-0141287.

PR	01-JUL-1999;	9905-01418484;
PR	02-JUL-1999;	9905-01421564;
PR	06-JUL-1999;	9905-01424055;
PR	08-JUL-1999;	9905-01424303;
PR	09-JUL-1999;	9905-01424803;
PR	12-JUL-1999;	9905-01424920;
PR	13-JUL-1999;	9905-01435422;
PR	14-JUL-1999;	9905-01436244;
PR	15-JUL-1999;	9905-01440055;
PR	16-JUL-1999;	9905-01440085;
PR	16-JUL-1999;	9905-01440086;
PR	19-JUL-1999;	9905-01443334;
PR	19-JUL-1999;	9905-01443343;
PR	19-JUL-1999;	9905-01443325;
PR	19-JUL-1999;	9905-01443321;
PR	19-JUL-1999;	9905-01443312;
PR	21-JUL-1999;	9905-01444814;
PR	21-JUL-1999;	9905-01445086;
PR	21-JUL-1999;	9905-01450086;
PR	22-JUL-1999;	9905-01450855;
PR	22-JUL-1999;	9905-01450879;
PR	22-JUL-1999;	9905-01450899;
PR	22-JUL-1999;	9905-01451195;
PR	23-JUL-1999;	9905-01451445;
PR	23-JUL-1999;	9905-01451824;
PR	23-JUL-1999;	9905-01452218;
PR	26-JUL-1999;	9905-01453276;
PR	27-JUL-1999;	9905-01453913;
PR	27-JUL-1999;	9905-01454918;
PR	27-JUL-1999;	9905-01454919;
PR	28-JUL-1999;	9905-01454951;
PR	02-AUG-1999;	9905-01463866;
PR	02-AUG-1999;	9905-01463888;
PR	02-AUG-1999;	9905-01464330;
PR	03-AUG-1999;	9905-01470308;
PR	03-AUG-1999;	9905-01472004;
PR	04-AUG-1999;	9905-01473022;
PR	05-AUG-1999;	9905-01471922;
PR	05-AUG-1999;	9905-01472600;
PR	06-AUG-1999;	9905-01474303;
PR	06-AUG-1999;	9905-01474316;
PR	09-AUG-1999;	9905-01474793;
PR	09-AUG-1999;	9905-01479355;
PR	10-AUG-1999;	9905-01481711;
PR	11-AUG-1999;	9905-01483319;
PR	12-AUG-1999;	9905-01483341;
PR	13-AUG-1999;	9905-01486565;
PR	13-AUG-1999;	9905-01486864;
PR	16-AUG-1999;	9905-01493368;
PR	17-AUG-1999;	9905-01491715;
PR	18-AUG-1999;	9905-01494326;
PR	20-AUG-1999;	9905-01493723;
PR	20-AUG-1999;	9905-01497223;
PR	20-AUG-1999;	9905-01498929;
PR	23-AUG-1999;	9905-01499022;
PR	23-AUG-1999;	9905-01499302;
PR	25-AUG-1999;	9905-01505884;
PR	26-AUG-1999;	9905-01505864;
PR	27-AUG-1999;	9905-01510665;
PR	27-AUG-1999;	9905-01510666;
PR	27-AUG-1999;	9905-01510680;
PR	30-AUG-1999;	9905-01515303;
PR	31-AUG-1999;	9905-01514338;
PR	01-SEP-1999;	9905-01519330;
PR	07-SEP-1999;	9905-01519363;
PR	10-SEP-1999;	9905-01530700;
PR	13-SEP-1999;	9905-01530718;
PR	15-SEP-1999;	9905-01545018;

PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

Search completed: July 24, 2003, 11:53:41
Job time : 14.0015 secs

Query Match	26.0%	Score 56	DB 21	Length 219
Best Local Similarity	33.3%	Pred. No. 23		
Matches	10	Conservative	8	Mismatches 12; Indels 0; Gaps 0;
OY	9	PSRRKEWMENAGNKIYTMADKTKISKIMTE	38	
		↓ : : : : :		
db	105	PSLQKQWYSVLGNDDYRGNEAQLSKVLTO	134	

PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.

GenCore version 5.1.6.
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using BW model

Run on: July 24, 2003, 11:53:53 ; Search time 3.57247 Seconds
(without alignments)
473.743 Million cell updates/sec

Title: US-09-991-681-30

Perfect score: 215
Sequence: 1 SPKYEKKDPKSRKEMWENAGNIYMAADKTISKLMTEYK 40

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, AA.*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	215	100.0	40	US-09-065-383-30	Sequence 30, Appl
2	215	100.0	518	US-09-065-383-27	Sequence 27, Appl
3	57	26.5	275	US-09-328-352-5591	Sequence 5591, Ap
4	55.5	25.8	244	US-09-107-532A-3889	Sequence 3889, Ap
5	55	25.6	351	US-09-431-976-2	Sequence 2, Appl
6	55	25.6	351	US-09-021-286-2	Sequence 2, Appl
7	52.5	24.4	234	US-09-296-754-2	Sequence 2, Appl
8	51.5	24.0	489	US-09-986-536-2	Sequence 2, Appl
9	51	23.7	528	US-08-793-229-35	Sequence 35, Appl
10	51	23.7	528	US-09-285-957-35	Sequence 35, Appl
11	51	23.7	1729	US-09-553-690-2	Sequence 2, Appl
12	51	23.7	4544	US-08-469-486-52	Sequence 52, Appl
13	51	23.7	4544	US-08-469-658-52	Sequence 52, Appl
14	50.5	23.5	489	US-09-545-814-29	Sequence 29, Appl
15	50.5	23.5	631	US-09-107-532A-6640	Sequence 6640, Ap
16	50.5	23.5	995	US-08-673-789-5	Sequence 5, Appl
17	50	23.3	92	US-08-713-939A-79	Sequence 79, Appl
18	50	23.3	92	US-09-036-578-79	Sequence 79, Appl
19	50	23.3	92	US-09-550-374-79	Sequence 79, Appl
20	50	23.3	92	US-09-943-906-79	Sequence 79, Appl
21	50	23.3	291	US-08-568-459A-13	Sequence 13, Appl
22	50	23.3	291	US-08-487-826B-25	Sequence 25, Appl
23	50	23.3	291	US-09-210-288-13	Sequence 13, Appl
24	50	23.3	397	US-08-774-104A-2	Sequence 2, Appl
25	50	23.3	1115	US-08-568-459A-2	Sequence 2, Appl
26	50	23.3	1115	US-08-487-826B-2	Sequence 2, Appl
27	50	23.3	1115	US-09-210-288-2	Sequence 2, Appl

28	50	23.3	1115	6	5198347-6	Patent No. 5198347
29	49.5	23.0	147	4	US-09-252-991A-23806	Sequence 23806, A
30	49	22.8	220	4	US-09-252-991A-22823	Sequence 22823, A
31	49	22.8	476	4	US-09-346-408-12	Sequence 12, Appl
32	49	22.8	970	2	US-08-673-789-7	Sequence 7, Appl
33	49	22.8	973	1	US-08-162-809-8	Sequence 8, Appl
34	49	22.8	973	1	US-08-162-809-10	Sequence 10, Appl
35	49	22.8	988	1	US-08-162-809-14	Sequence 14, Appl
36	49	22.8	995	1	US-08-162-809-18	Sequence 18, Appl
37	49	22.8	1011	1	US-08-162-809-12	Sequence 12, Appl
38	48.5	22.6	261	1	US-08-624-125-18	Sequence 18, Appl
39	48.5	22.6	261	1	US-08-937-155-18	Sequence 18, Appl
40	48.5	22.6	569	1	US-08-306-231-3	Sequence 3, Appl
41	48.5	22.6	672	3	US-09-040-843-4	Sequence 4, Appl
42	48.5	22.6	672	4	US-09-621-855-4	Sequence 4, Appl
43	48.5	22.6	866	3	US-09-040-843-2	Sequence 2, Appl
44	48.5	22.6	866	4	US-09-621-855-2	Sequence 2, Appl
45	48.5	22.6	896	2	US-08-640-389A-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-09-065-383-30
; Sequence 30, Application US/09065383
; Patent No. 6391543
; GENERAL INFORMATION:
; APPLICANT: BILLING-MEDEL, PATRICIA
; APPLICANT: COHEN, MAURICE
; APPLICANT: COLPITTS, TRACEY L.
; APPLICANT: FRIEDMAN, PAULA N.
; APPLICANT: GORDON, JULIAN
; APPLICANT: GRANADOS, EDWARD N.
; APPLICANT: HODGES, STEVEN C.
; APPLICANT: KLASS, MICHAEL R.
; APPLICANT: KRATOCHVIL, JON D.
; APPLICANT: ROBERTS-RAPP, LISA
; APPLICANT: RUSSELL, JOHN C.
; APPLICANT: STROUPE, STEPHEN D.
; TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
; TITLE OF INVENTION: FOR DETECTING DISEASES OF THE PROSTATE
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastrSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/065,383
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/842,385
; FILING DATE: 23-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Becker, Cheryl L.
; REGISTRATION NUMBER: 35,441
; REFERENCE/DOCKET NUMBER: 6084, US, P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847/935-1729
; TELEFAX: 847/938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6391543e
US-09-065-383-30

Query Match 100.0%; Score 215; DB 4; Length 40;
Best Local Similarity 100.0%; Pred. No. 5.6e-23;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPKVEKQPSRKKEWENAGNKIYTMADKTIISKLTETK 40
Db 1 SPKVEKQPSRKKEWENAGNKIYTMADKTIISKLTETK 40

RESULT 2
US-09-065-383-27

Sequence 27, Application US/09065383
Patent No. 6391543

GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA

APPLICANT: COHEN, MAURICE
APPLICANT: COLPITS, TRACEY L.

APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GORDON, JULIAN

APPLICANT: GRANADOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.

APPLICANT: KLAAS, MICHAEL R.
APPLICANT: KRATOCHVIL, JON D.

APPLICANT: ROBERTS-RAPP, LISA
APPLICANT: RUSSELL, JOHN C.

APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL

TITLE OF INVENTION: FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories

STREET: 100 Abbott Park Road
CITY: Abbott Park

STATE: IL
COUNTRY: USA

ZIP: 60064-3500
COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS
SOFTWARE: FastsEQ for Windows Version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/065,383

FILING DATE:
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/842,385

FILING DATE: 23-APR-1997
ATTORNEY/AGENT INFORMATION:

NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441

REFERENCE/DOCKET NUMBER: 6084.US.PI
TELECOMMUNICATION INFORMATION:

TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623

TELEX:
INFORMATION FOR SEQ ID NO: 27:

SEQUENCE CHARACTERISTICS:
LENGTH: 518 amino acids

TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: No. 6391543e

US-09-065-383-27

Query Match 100.0%; Score 215; DB 4; Length 518;
Best Local Similarity 100.0%; Pred. No. 1.3e-21;

Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPKVEKQPSRKKEWENAGNKIYTMADKTIISKLTETK 40
Db 332 SPKVEKQPSRKKEWENAGNKIYTMADKTIISKLTETK 371

RESULT 3
US-09-328-352-5591

Sequence 5591, Application US/09328352
Patent No. 6562958

GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
TITLE OF INVENTION: BAUMANNIT FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352

CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252

SEQ ID NO 5591
LENGTH: 275

TYPE: PRT
ORGANISM: Acinetobacter baumannii

US-09-328-352-5591

Query Match 26.5%; Score 57; DB 4; Length 275;
Best Local Similarity 45.5%; Pred. No. 4.1;
Matches 10; Conservative 4; Mismatches 6; Indels 2; Gaps 1;

Qy 2 PKVEKQPSRKKEWENAGNKI 23
Db 246 PVTKKEPKM--WQKQKGMKL 265

RESULT 4
US-09-107-532A-3889

Sequence 3889, Application US/09107532A
Patent No. 6583275

GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:

ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street

CITY: Waltham
STATE: Massachusetts

COUNTRY: USA
ZIP: 02154

COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660

COMPUTER: PC
OPERATING SYSTEM: <Unknown>

SOFTWARE: ASCII
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598

FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571

FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:

NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489

REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:

TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 3889:
SEQUENCE CHARACTERISTICS:

LENGTH: 244 amino acids

```

;
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...244
; SEQUENCE DESCRIPTION: SEQ ID NO: 3889:
US-09-107-532A-3889

Query Match      25.8%; Score 55.5; DB 4; Length 244;
Best Local Similarity 31.8%; Pred. No. 5.7;
Matches 14; Conservative 7; Mismatches 16; Indels 7; Gaps 1;

Oy      3 KVEKKDPSRKKEWENAGNKIYT-----MAADTKISKLMTEY 39
Db      59 KLVSYDKDKAKRYWKAKKELGVDSLRFDLMASSDDSSKVIYEV 102

RESULT 5
US-09-431-976-2
; Sequence 2, Application US/09431976
; Patent No. 6423520
; GENERAL INFORMATION:
; APPLICANT: Konkling, Mark
; APPLICANT: Mendu, Nandini
; APPLICANT: Song, Wen
; TITLE OF INVENTION: Regulation of Quinolinate Phosphoribosyl Transferase
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 5051-338
; CURRENT APPLICATION NUMBER: US/09/431,976
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: 09/021,286
; PRIOR FILING DATE: 1998-02-10
; PRIOR APPLICATION NUMBER: 60/049,471
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Nicotiana tabacum
US-09-431-976-2

Query Match      25.6%; Score 55; DB 4; Length 351;
Best Local Similarity 52.9%; Pred. No. 10;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Oy      8 DPSRKKEWENAGNKIY 24
Db      106 DPLSKVEWYVNDGDKVH 122

RESULT 6
US-09-021-286-2
; Sequence 2, Application US/09021286B
; Patent No. 6586661
; GENERAL INFORMATION:
; APPLICANT: Konkling, Mark
; APPLICANT: Mendu, Nandini
; APPLICANT: Song, Wen
; TITLE OF INVENTION: Regulation of Quinolinate Phosphoribosyl Transferase
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 5051-338
; CURRENT APPLICATION NUMBER: US/09/021,286B
; PRIOR FILING DATE: 1998-02-10
; PRIOR APPLICATION NUMBER: 60/049,471
; EARLIER FILING DATE: 1997-06-12
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
```

```

; LENGTH: 351
; TYPE: PRT
; ORGANISM: Nicotiana tabacum
US-09-021-286-2

Query Match      25.6%; Score 55; DB 4; Length 351;
Best Local Similarity 52.9%; Pred. No. 10;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Oy      8 DPSRKKEWENAGNKIY 24
Db      106 DPLSKVEWYVNDGDKVH 122

RESULT 7
US-09-296-754-2
; Sequence 2, Application US/09296754A
; Patent No. 6316695
; GENERAL INFORMATION:
; APPLICANT: Korea Kumho Petrochemical Co. Ltd.
; TITLE OF INVENTION: Isopentenyl diphosphate isomerase from Hevea Brasiliensis
; FILE REFERENCE: PX99102/US
; CURRENT APPLICATION NUMBER: US/09/296,754A
; PRIOR FILING DATE: 1999-04-22
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: KOPATIN 1.0
; SEQ ID NO 2
; LENGTH: 234
; TYPE: PRT
; ORGANISM: IPHb (IPP isomerase from Hevea Brasiliensis)
US-09-296-754-2

Query Match      24.4%; Score 52.5; DB 4; Length 234;
Best Local Similarity 52.2%; Pred. No. 14;
Matches 12; Conservative 1; Mismatches 9; Indels 1; Gaps 1;

Oy      14 EWMENAGNKIYTMAD--KTISKL 35
Db      211 KWMENVNGTKLKVNDKTIHKL 233

RESULT 8
US-09-986-536-2
; Sequence 2, Application US/09986536
; Patent No. 6461841
; GENERAL INFORMATION:
; APPLICANT: GEDEKE, BIRGIT
; APPLICANT: HUMMEL, WERNER
; APPLICANT: BOMMARIUS, ANDREAS
; TITLE OF INVENTION: L-Amino Acid Oxidase from Rhodococcus Species
; FILE REFERENCE: 215209USOX
; CURRENT APPLICATION NUMBER: US/09/986,536
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: DE 100 55 512.8
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Rhodococcus opacus
US-09-986-536-2

Query Match      24.0%; Score 51.5; DB 4; Length 489;
Best Local Similarity 35.6%; Pred. No. 47;
Matches 16; Conservative 6; Mismatches 14; Indels 9; Gaps 3;

Oy      3 KVEKKDPSRK-----KEMWENAGNKIYTMAA--DKTISKLMTEY 39
Db      313 KAAKPSGSKLGIEYSRRWMTB-DRIYCGASNTDKDISQIMFYY 356

RESULT 9
```

US-08-793-229-35
; Sequence 35, Application US/08793229
; Patent No. 5891703
; GENERAL INFORMATION:
; APPLICANT: VAN DER LAAN, Jan Metake
; APPLICANT: RIEMENS, Adriana Martina
; APPLICANT: QUAX, Wilhelmus Johannes
; TITLE OF INVENTION: Mutated Penicillin G Acylase Genes
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boenhen Hulbert & Berghoff
; STREET: 300 South Wacker Drive
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/793,229
; FILING DATE: 23-APR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/03249
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; REFERENCE/DOCKET NUMBER: 97075
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312)913-0001
; TELEFAX: (312)913-0002
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 528 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-793-229-35
Query Match 23.7%; Score 51; DB 2; Length 528;
Best Local Similarity 27.7%; Pred. No. 61;
Matches 13; Conservative 3; Mismatches 11; Indels 20; Gaps 1;
QY 14 EWMENAGNKIY-----TMAADKTSKLTMEYK 40
DB 355 EWMNNLHDKLFMDLGDGYGITEITDHRYGASLAYKNISKSTNYK 401
RESULT 10
US-09-285-957-35
; Sequence 35, Application US/09285957
; Patent No. 6033823
; GENERAL INFORMATION:
; APPLICANT: VAN DER LAAN, Jan Metake
; APPLICANT: RIEMENS, Adriana Martina
; APPLICANT: QUAX, Wilhelmus Johannes
; TITLE OF INVENTION: Mutated Penicillin G Acylase Genes
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boenhen Hulbert & Berghoff
; STREET: 300 South Wacker Drive
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/285,957
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/793,229
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; REFERENCE/DOCKET NUMBER: 97075
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312)913-0001
; TELEFAX: (312)913-0002
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 528 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-285-957-35
Query Match 23.7%; Score 51; DB 3; Length 528;
Best Local Similarity 27.7%; Pred. No. 61;
Matches 13; Conservative 3; Mismatches 11; Indels 20; Gaps 1;
QY 14 EWMENAGNKIY-----TMAADKTSKLTMEYK 40
DB 355 EWMNNLHDKLFMDLGDGYGITEITDHRYGASLAYKNISKSTNYK 401
RESULT 11
US-09-553-690-2
; Sequence 2, Application US/09553690
; Patent No. 6476296
; GENERAL INFORMATION:
; APPLICANT: Fischer, Robert L.
; APPLICANT: Choi, Yeonhee
; APPLICANT: Hannon, Mike
; TITLE OF INVENTION: The Regents of the University of California
; TITLE OF INVENTION: Nucleic Acids That Control Seed and
; FILE REFERENCE: 023070-099900US
; CURRENT APPLICATION NUMBER: US/09/553,690
; CURRENT FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1729
; TYPE: PRT
; ORGANISM: Arabidopsis sp.
; FEATURE:
; OTHER INFORMATION: ATROPOS (ATR) amino acid sequence
US-09-553-690-2
Query Match 23.7%; Score 51; DB 4; Length 1729;
Best Local Similarity 32.3%; Pred. No. 2.6e+02;
Matches 10; Conservative 10; Mismatches 9; Indels 2; Gaps 2;
QY 5 EKDPSRKKEWMENAGNKIYTMAADKTSKLT 35
DB 722 EKKD-KKEKWEKE-RVFRGRADSFIAFM 750
RESULT 12
US-08-469-486-52
; Sequence 52, Application US/08469486
; Patent No. 5739281
; GENERAL INFORMATION:
; APPLICANT: Thoegeersen, Hans Christian
; APPLICANT: Holteit, Thor Las
; APPLICANT: Ezerodt, Michael
; TITLE OF INVENTION: Improved method for the refolding of
; TITLE OF INVENTION: proteins

NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version
SOFTWARE: #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,486
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/192,060
FILING DATE: February 4, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Paul T. Clark
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 06363/002001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617 542 5070
TELEFAX: 617 542 8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:
LENGTH: 4544 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-469-486-52

Query Match 23.7%; Score 51; DB 1; Length 4544;
Best Local Similarity 36.4%; Pred. No. 8.3e+02;
Matches 12; Conservative 6; Mismatches 13; Indels 2; Gaps 1;

Qy 8 DPSRKKEWENAGN-KIYMAADKTSKLMTE 38
Db 4009 DPLRGTYWSDMGNHPIETAMDGTLRETLVQ 4041

RESULT 13
US-08-469-658-52
Sequence 52, Application US/08469658
Patent No. 5917018
GENERAL INFORMATION:
APPLICANT: Th egeeren, Hans Christian
APPLICANT: Hollet, Thor Las
APPLICANT: Etzerodt, Michael
TITLE OF INVENTION: IMPROVED METHOD FOR THE REPODLING OF
TITLE OF INVENTION: PROTEINS
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version
SOFTWARE: #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,658

FILING DATE: June 5, 1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/192,060
FILING DATE: February 4, 1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Paul T. Clark
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 06363/002002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617 542 5070
TELEFAX: 617 542 8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:
LENGTH: 4544 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-469-658-52

Query Match 23.7%; Score 51; DB 2; Length 4544;
Best Local Similarity 36.4%; Pred. No. 8.3e+02;
Matches 12; Conservative 6; Mismatches 13; Indels 2; Gaps 1;

Qy 8 DPSRKKEWENAGN-KIYMAADKTSKLMTE 38
Db 4009 DPLRGTYWSDMGNHPIETAMDGTLRETLVQ 4041

RESULT 14
US-09-545-814-29
Sequence 29, Application US/09545814
Patent No. 6416977
GENERAL INFORMATION:
APPLICANT: Becher, Anna M.
TITLE OF INVENTION: FLEA CHITTINASE NUCLEIC ACID MOLECULES, PROTEINS AND
TITLE OF INVENTION: USES THEREOF
FILE REFERENCE: FC-5-CI
CURRENT FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: 60/128,833
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 40
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 29
LENGTH: 489
TYPE: PRT
ORGANISM: Tagged Ctenocephalides felis
US-09-545-814-29

Query Match 23.5%; Score 50.5; DB 4; Length 489;
Best Local Similarity 29.8%; Pred. No. 65;
Matches 14; Conservative 6; Mismatches 16; Indels 11; Gaps 2;

Qy 5 EKDPSSKK--EWMENAGNKIYMAADK-----TISKLMTEYK 40
Db 29 DDKDPVKLIQIAGMGAGGKRYSTVAEKRSAPFIRSVDPFMEYK 75

RESULT 15
US-09-107-532A-6640
Sequence 6640, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucetice-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
TITLE OF INVENTION: ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION

```
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051,571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 6640:
SEQUENCE CHARACTERISTICS:
LENGTH: 631 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...631
SEQUENCE DESCRIPTION: SEQ ID NO: 6640:
US-09-107-532A-6640

Query Match      23.5%; Score 50.5; DB 4; Length 631;
Best Local Similarity 29.5%; Pred. No. 88;
Matches 13; Conservative 10; Mismatches 12; Indels 9; Gaps 2;

QY      2 PRVEKKDPSRKKEW-----WENAGNKITYTMADKTIKSKMTE 38
DB      549 PRIQSKPEPAKKKLSYMEQKEWTIEDRIAEI--EEKISLLOEE 590
```

Search completed: July 24, 2003, 12:16:24
Job time : 4.57247 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 11:57:13 ; Search time 4.62665 Seconds
(without alignments)
1026.746 Million cell updates/sec

Title: US-09-991-681-30

Perfect score: 215
Sequence: 1 SPKYEKKDPSRRKEMWENAGNKIYTMADKTIKLMTEYK 40

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA:*

1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*\n2: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep:*\n3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*\n4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep:*\n5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*\n6: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUBCOMB.pep:*\n7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*\n8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep:*\n9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*\n10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*\n11: /cgn2_6/ptodata/2/pubppaa/US09C_PUBCOMB.pep:*\n12: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep:*\n13: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*\n14: /cgn2_6/ptodata/2/pubppaa/US10B_PUBCOMB.pep:*\n15: /cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pep:*\n16: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep:*\n17: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*\n18: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	59.5	27.7	388	9 US-09-815-242-11623	Sequence 11623, A
2	56.5	26.3	388	9 US-09-815-242-11458	Sequence 11458, A
3	56.5	26.3	388	10 US-09-881-7524-350	Sequence 350, App
4	55	25.6	351	10 US-09-963-340-2	Sequence 2, Appl
5	54	25.1	566	15 US-10-156-761-9082	Sequence 9082, App
6	53	24.7	462	10 US-09-712-363-709	Sequence 209, App
7	52.5	24.4	431	15 US-10-050-704-106	Sequence 106, App
8	52	24.2	336	15 US-10-160-293-4	Sequence 4, Appl
9	52	24.2	396	11 US-10-160-293-2	Sequence 2, Appl
10	52	24.2	468	11 US-09-931-836-67	Sequence 67, Appl
11	52	24.2	468	13 US-10-036-342-67	Sequence 67, Appl
12	52	24.2	468	14 US-10-006-867-150	Sequence 150, App
13	52	24.2	468	14 US-10-077-040-1	Sequence 1, Appl
14	52	24.2	468	14 US-10-063-547-150	Sequence 150, App
15	52	24.2	468	14 US-10-036-041-67	Sequence 67, Appl

16	52	24.2	468	15 US-10-028-072-90	Sequence 90, Appl
17	52	24.2	468	15 US-10-035-855-67	Sequence 67, Appl
18	52	24.2	468	15 US-10-063-616-150	Sequence 150, Appl
19	52	24.2	468	15 US-10-121-049-90	Sequence 90, Appl
20	52	24.2	468	15 US-10-123-904-90	Sequence 90, Appl
21	52	24.2	468	15 US-10-140-470-90	Sequence 90, Appl
22	52	24.2	468	15 US-10-063-502-150	Sequence 150, App
23	52	24.2	468	15 US-10-175-146-90	Sequence 90, Appl
24	52	24.2	468	15 US-10-176-918-90	Sequence 90, Appl
25	52	24.2	468	15 US-10-176-821-90	Sequence 90, Appl
26	52	24.2	468	15 US-10-036-214-67	Sequence 67, Appl
27	52	24.2	468	15 US-10-137-865-90	Sequence 90, Appl
28	52	24.2	468	15 US-10-140-474-90	Sequence 90, Appl
29	52	24.2	468	15 US-10-035-419-67	Sequence 67, Appl
30	52	24.2	468	15 US-10-142-431-90	Sequence 90, Appl
31	52	24.2	468	15 US-10-143-114-90	Sequence 90, Appl
32	52	24.2	468	15 US-10-140-002-90	Sequence 90, Appl
33	52	24.2	468	15 US-10-036-160-67	Sequence 67, Appl
34	52	24.2	468	15 US-10-142-419-90	Sequence 90, Appl
35	52	24.2	468	15 US-10-035-958-67	Sequence 67, Appl
36	52	24.2	468	15 US-10-036-150-67	Sequence 67, Appl
37	52	24.2	468	15 US-10-063-518-150	Sequence 150, App
38	52	24.2	468	15 US-10-123-262-90	Sequence 90, Appl
39	52	24.2	468	15 US-10-142-423-90	Sequence 90, Appl
40	52	24.2	468	15 US-10-063-598-150	Sequence 150, App
41	52	24.2	468	15 US-10-227-693-150	Sequence 150, App
42	52	24.2	468	15 US-10-121-050-90	Sequence 90, Appl
43	52	24.2	468	15 US-10-141-755-90	Sequence 90, Appl
44	52	24.2	468	15 US-10-143-032-90	Sequence 90, Appl
45	52	24.2	468	15 US-10-123-108-90	Sequence 90, Appl

ALIGNMENTS

RESULT 1
US-09-815-242-11623
Sequence 11623, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA 011A
CURRENT APPLICATION NUMBER: US/09/815,242
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11623
LENGTH: 388
TYPE: PRT
ORGANISM: Helicobacter pylori

US-09-815-242-11623

Query Match

Best Local Similarity 27.7%; Score 59.5; DB 9; Length 388;
Matches 13; Conservative 4; Mismatches 13; Indels 5; Gaps 1;

QY 11 RKKEWMENAGNKIYTMADK-----TISKLMTEYK 40

DB 217 RLKLMWSNLQNSLFTLLPRLANLRLISDLPESYQ 251

RESULT 2

US-09-815-242-11458
Sequence 11458, Application US/09815242
Patent No. US2002061569A1
GENERAL INFORMATION:

APPLICANT: Haseibeck, Robert

APPLICANT: Ohlsen, Kari L.

APPLICANT: Zyskind, Judith W.

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John D.

APPLICANT: Carr, Grant J.

APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in

Prokaryotes

FILE REFERENCE: ELITRA.011A

CURRENT APPLICATION NUMBER: US/09/815,242

CURRENT FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 11458

LENGTH: 388

TYPE: PR

ORGANISM: Helicobacter pylori

FEATURE:

NAME/KEY: VARIANT

LOCATION: (1)...(388)

OTHER INFORMATION: Xaa = Any Amino Acid

US-09-815-242-11458

Query Match

Best Local Similarity 26.3%; Score 56.5; DB 9; Length 388;
Matches 12; Conservative 5; Mismatches 13; Indels 5; Gaps 1;

QY 11 RKKEWMENAGNKIYTMADK-----TISKLMTEYK 40

DB 217 RLKLMWSNLQNSLFTLLPRLANLRLISDLPESYQ 251

RESULT 3

US-09-881-752A-350

Sequence 350, Application US/09881752A

Patent No. US20020115078A1

GENERAL INFORMATION:

APPLICANT: Kleantous, Harold

APPLICANT: Al-Garawi, Amal

APPLICANT: Miller, Charles

APPLICANT: Tomb, Jean-Francois

APPLICANT: Comen, Raymond P

TITLE OF INVENTION: Identification of Polynucleotides

TITLE OF INVENTION: Encoding No. US20020115078A1 Helicobacter Polypeptides in the

FILE REFERENCE: 06132/041002

CURRENT APPLICATION NUMBER: US/09/881,752A

CURRENT FILING DATE: 2001-06-15

PRIOR APPLICATION NUMBER: US 08/833,457

PRIOR FILING DATE: 1997-04-01

NUMBER OF SEQ ID NOS: 370

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 350

LENGTH: 388

TYPE: PR

ORGANISM: Helicobacter pylori

US-09-881-752A-350

Query Match

Best Local Similarity 26.3%; Score 56.5; DB 10; Length 388;
Matches 12; Conservative 5; Mismatches 13; Indels 5; Gaps 1;

QY 11 RKKEWMENAGNKIYTMADK-----TISKLMTEYK 40

DB 217 RLKLMWSNLQNSLFTLLPRLANLRLISDLPESYQ 251

RESULT 4

US-09-963-340-2

Sequence 2, Application US/09963340

Patent No. US2002010815A1

GENERAL INFORMATION:

APPLICANT: Conkling, Mark

APPLICANT: Mendu, Nandini

APPLICANT: Song, Wen

TITLE OF INVENTION: Regulation of Quinolinate Phosphoribosyl Transferase

FILE REFERENCE: 5051-338

CURRENT APPLICATION NUMBER: US/09/963,340

CURRENT FILING DATE: 2001-09-24

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/021,286

PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-10

NUMBER OF SEQ ID NOS: 3

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 2

LENGTH: 351

TYPE: PR

ORGANISM: Nicotiana tabacum

US-09-963-340-2

Query Match

Best Local Similarity 25.6%; Score 55; DB 10; Length 351;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 8 DPKKEWMENAGNKIY 24

DB 106 DPKKEWMENAGNKIY 122

RESULT 5

US-10-156-761-9082

Sequence 9082, Application US/10156761

Publication No. US20030119018A1

GENERAL INFORMATION:

APPLICANT: OMURA, SATOSHI

APPLICANT: IKEDA, HARUO

APPLICANT: ISHIKAWA, JUN

APPLICANT: HORIKAWA, HIROSHI

APPLICANT: SHIBA, TADAYOSHI

APPLICANT: SAKAKI, YOSHIYUKI

APPLICANT: HATTORI, MASAHIRA

TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

FILE REFERENCE: 249-262

CURRENT APPLICATION NUMBER: US/10/156,761


```

; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9082
; LENGTH: 566
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9082

```

```

Query Match      25.1%; Score 54; DB 15; Length 566;
Best Local Similarity 42.1%; Pred. No. 59;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

```

```

Qy      4 VEKQPSRKKEWENAGNK 22
Db      494 VKEQNPFRSKWMIWVGTK 512

```

```

RESULT 6
US-09-712-363-209
; Sequence 209, Application US/09712363
; Patent No. US20020164588A1
; GENERAL INFORMATION:
; APPLICANT: Eisenberg, David
; APPLICANT: Marcotte, Edward M.
; APPLICANT: Roetsch, David M.
; TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
; FILE REFERENCE: 07419-032001
; CURRENT FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: US/09/712,363
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,531
; PRIOR FILING DATE: 2000-02-01
; PRIOR APPLICATION NUMBER: 60/117,844
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: 60/118,206
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: 60/126,593
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 60/134,093
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/134,092
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/165,124
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/165,086
; PRIOR FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 292
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 209
; LENGTH: 462
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-712-363-209

```

```

Query Match      24.7%; Score 53; DB 10; Length 462;
Best Local Similarity 31.6%; Pred. No. 64;
Matches 12; Conservative 8; Mismatches 16; Indels 2; Gaps 1;

```

```

Qy      1 SPVKEKPSRKKEWENAGNKIYMAADKTSKLMTE 38
Db      179 APYAGDVVFR--AWMDLAGNRRAGPPIARAVSKVIAE 214

```

```

RESULT 7
US-10-050-704-106
; Sequence 106, Application US/10050704
; Publication No. US20030050442A1

```

```

; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 62 Human Secreted Proteins
; FILE REFERENCE: P2039P1
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US/10/050,704
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: 09/684,524
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/US00/08979
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 60/130,991
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 344
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 106
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-050-704-106

```

```

Query Match      24.4%; Score 52.5; DB 15; Length 231;
Best Local Similarity 30.0%; Pred. No. 36;
Matches 12; Conservative 3; Mismatches 12; Indels 13; Gaps 1;

```

```

Qy      4 VEKQPSRK-----KEMWENAGNKIYMAADK 30
Db      182 IRKQPSPLANTTYNIFIMDKTWMHNSKNTKLAKK 221

```

```

RESULT 8
US-10-160-293-4
; Sequence 4, Application US/10160293
; Publication No. US20030022208A1
; GENERAL INFORMATION:
; APPLICANT: Lindquist, Erika et al.
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; FILE REFERENCE: CU001241-PROV
; CURRENT FILING DATE: 2002-06-04
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Human
US-10-160-293-4

```

```

Query Match      24.2%; Score 52; DB 15; Length 336;
Best Local Similarity 32.6%; Pred. No. 62;
Matches 15; Conservative 8; Mismatches 9; Indels 14; Gaps 3;

```

```

Qy      6 KQDPSRKKEWEN-AG--NKIYMAA-----DKITSKLMTE 37
Db      77 EEDNKPRTSTENQAKIPEKVTMAAIQDGLANGENDVTSNTLT 122

```

```

RESULT 9
US-10-160-293-2
; Sequence 2, Application US/10160293
; Publication No. US20030022208A1
; GENERAL INFORMATION:
; APPLICANT: Lindquist, Erika et al.
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; FILE REFERENCE: CU001241-PROV
; CURRENT FILING DATE: 2002-06-04
; NUMBER OF SEQ ID NOS: 5

```

```
SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Human
US-10-160-293-2

Query Match      24.2%; Score 52; DB 15; Length 396;
Best Local Similarity 32.6%; Pred. No. 74;
Matches 15; Conservative 8; Mismatches 9; Indels 14; Gaps 3;

QY      6 KQDPRKKKEMEN-AG---NKITYTMA-----DKTISKLMT 37
Db      137 EEPNKPSTSWTENQAKIKPKVTPMAIODGLAKGENDETNSNTLT 182

RESULT 10
US-09-931-836-67
; Sequence 67, Application US/09931836
; Publication No. US20030027249A1
; GENERAL INFORMATION:
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gueney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3030R1C1
; CURRENT FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/112514
; PRIOR FILING DATE: 1998-12-15
; PRIOR APPLICATION NUMBER: 60/113300
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 60/113430
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/113605
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/113621
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/114140
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/115552
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/116843
; PRIOR FILING DATE: 1999-01-22
; PRIOR APPLICATION NUMBER: 60/125774
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 60/125778
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 60/125826
; PRIOR FILING DATE: 1999-03-24
; PRIOR APPLICATION NUMBER: 60/127035
; PRIOR FILING DATE: 1999-03-31
; PRIOR APPLICATION NUMBER: 60/127706
; PRIOR FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: 60/129122
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 60/130359
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/131270
; PRIOR FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: 60/131272
; PRIOR FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: 60/131291
```

```

; PRIOR FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: 60/132371
; PRIOR FILING DATE: 1999-05-04
; PRIOR APPLICATION NUMBER: 60/133379
; PRIOR FILING DATE: 1999-05-04
; PRIOR APPLICATION NUMBER: 60/133383
; PRIOR FILING DATE: 1999-05-04
; PRIOR APPLICATION NUMBER: 60/135750
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 60/138166
; PRIOR FILING DATE: 1999-06-08
; PRIOR APPLICATION NUMBER: 60/144791
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/146970
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: 60/162506
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: 09/311832
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 09/380142
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: 09/644848
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 09/747259
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 09/816744
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 09/854208
; PRIOR FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: 09/854280
; PRIOR FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: 09/874503
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: 09/869599
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 09/908,827
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: PCT/US99/10733
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: PCT/US99/28551
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30720
; PRIOR FILING DATE: 1999-12-22
; PRIOR APPLICATION NUMBER: PCT/US00/05601
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: PCT/US00/05841
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: PCT/US00/14042
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: PCT/US00/15264
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: PCT/US00/23522
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: PCT/US00/32678
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US00/34956
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: PCT/US01/06520
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: PCT/US01/17800
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: PCT/US01/19692
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: PCT/US01/21066
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: PCT/US01/21735
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 80
; SEQ ID NO 67
; LENGTH: 468
; TYPE: PRT
```

```

: ORGANISM: Homo Sapien
US-09-931-836-67

Query Match      24.2%   Score 52; DB 11; Length 468;
Best Local Similarity 32.6%; Pred. No. 88;
Matches 15; Conservative 8; Mismatches 9; Indels 14; Gaps 3

Oy      6 KCDPSRKKEWMEN-AG---NKITYTAA-----DKITSKMT 37
Db      209 EEDPKRPTSWTENQAKIPEKVTPMAAIQDGLAKGNDETVSNLT 254

RESULT 11
: Sequence 67, Application US/10036342
: Publication No. US20020090681A1
: GENERAL INFORMATION:
: APPLICANT: Desnoyers, Luc
: APPLICANT: Eaton, Dan L.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Pan, James
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Watanabe, Colin K.
: APPLICANT: Wood, William I.
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE REFERENCE: P3030R1CS
: CURRENT APPLICATION NUMBER: US/10/036.342
: CURRENT FILING DATE: 2001-12-26
: PRIOR APPLICATION NUMBER: 60/085579
: PRIOR FILING DATE: 1998-05-15
: PRIOR APPLICATION NUMBER: 60/112514
: PRIOR FILING DATE: 1998-12-15
: PRIOR APPLICATION NUMBER: 60/113300
: PRIOR FILING DATE: 1998-12-22
: PRIOR APPLICATION NUMBER: 60/113430
: PRIOR FILING DATE: 1998-12-23
: PRIOR APPLICATION NUMBER: 60/113605
: PRIOR FILING DATE: 1998-12-23
: PRIOR APPLICATION NUMBER: 60/113621
: PRIOR FILING DATE: 1998-12-23
: PRIOR APPLICATION NUMBER: 60/114140
: PRIOR FILING DATE: 1998-12-23
: PRIOR APPLICATION NUMBER: 60/115552
: PRIOR FILING DATE: 1999-01-12
: PRIOR APPLICATION NUMBER: 60/116843
: PRIOR FILING DATE: 1999-01-22
: PRIOR APPLICATION NUMBER: 60/125774
: PRIOR FILING DATE: 1999-03-23
: PRIOR APPLICATION NUMBER: 60/125778
: PRIOR FILING DATE: 1999-03-23
: PRIOR APPLICATION NUMBER: 60/125826
: PRIOR FILING DATE: 1999-03-24
: PRIOR APPLICATION NUMBER: 60/127035
: PRIOR FILING DATE: 1999-03-31
: PRIOR APPLICATION NUMBER: 60/127706
: PRIOR FILING DATE: 1999-04-05
: PRIOR APPLICATION NUMBER: 60/129122
: PRIOR FILING DATE: 1999-04-13
: PRIOR APPLICATION NUMBER: 60/130359
: PRIOR FILING DATE: 1999-04-21
: PRIOR APPLICATION NUMBER: 60/131270
: PRIOR FILING DATE: 1999-04-27
: PRIOR APPLICATION NUMBER: 60/131272
: PRIOR FILING DATE: 1999-04-27
: PRIOR APPLICATION NUMBER: 60/131291
: PRIOR FILING DATE: 1999-04-27
: PRIOR APPLICATION NUMBER: 60/132371
: PRIOR FILING DATE: 1999-05-04
: PRIOR APPLICATION NUMBER: 60/132379

```

```

? PRIOR FILING DATE: 1999-05-04
? PRIOR APPLICATION NUMBER: 60/13283
? PRIOR FILING DATE: 1999-05-04
? PRIOR APPLICATION NUMBER: 60/13750
? PRIOR FILING DATE: 1999-05-25
? PRIOR APPLICATION NUMBER: 60/138166
? PRIOR FILING DATE: 1999-06-08
? PRIOR APPLICATION NUMBER: 60/144791
? PRIOR FILING DATE: 1999-07-20
? PRIOR APPLICATION NUMBER: 60/146970
? PRIOR FILING DATE: 1999-08-03
? PRIOR APPLICATION NUMBER: 60/162506
? PRIOR FILING DATE: 1999-10-29
? PRIOR APPLICATION NUMBER: 09/311832
? PRIOR FILING DATE: 1999-05-14
? PRIOR APPLICATION NUMBER: 09/380142
? PRIOR FILING DATE: 1999-06-25
? PRIOR APPLICATION NUMBER: 09/644848
? PRIOR FILING DATE: 2000-08-22
? PRIOR APPLICATION NUMBER: 09/747259
? PRIOR FILING DATE: 2000-12-20
? PRIOR APPLICATION NUMBER: 09/816744
? PRIOR FILING DATE: 2001-03-22
? PRIOR APPLICATION NUMBER: 09/854208
? PRIOR FILING DATE: 2001-05-10
? PRIOR APPLICATION NUMBER: 09/854280
? PRIOR FILING DATE: 2001-05-10
? PRIOR APPLICATION NUMBER: 09/874503
? PRIOR FILING DATE: 2001-06-05
? PRIOR APPLICATION NUMBER: 09/869599
? PRIOR FILING DATE: 2001-06-29
? PRIOR APPLICATION NUMBER: 09/908.827
? PRIOR FILING DATE: 2001-07-18
? PRIOR APPLICATION NUMBER: PCT/US99/10733
? PRIOR FILING DATE: 1999-05-14
? PRIOR APPLICATION NUMBER: PCT/US99/28551
? PRIOR FILING DATE: 1999-12-02
? PRIOR APPLICATION NUMBER: PCT/US99/30720
? PRIOR FILING DATE: 1999-12-22
? PRIOR APPLICATION NUMBER: PCT/US00/05601
? PRIOR FILING DATE: 2000-03-01
? PRIOR APPLICATION NUMBER: PCT/US00/05841
? PRIOR FILING DATE: 2000-03-02
? PRIOR APPLICATION NUMBER: PCT/US00/14042
? PRIOR FILING DATE: 2000-05-22
? PRIOR APPLICATION NUMBER: PCT/US00/15264
? PRIOR FILING DATE: 2000-06-02
? PRIOR APPLICATION NUMBER: PCT/US00/23522
? PRIOR FILING DATE: 2000-08-23
? PRIOR APPLICATION NUMBER: PCT/US00/23328
? PRIOR FILING DATE: 2000-08-24
? PRIOR APPLICATION NUMBER: PCT/US00/32678
? PRIOR FILING DATE: 2000-12-01
? PRIOR APPLICATION NUMBER: PCT/US00/34956
? PRIOR FILING DATE: 2000-12-20
? PRIOR APPLICATION NUMBER: PCT/US01/06520
? PRIOR FILING DATE: 2001-02-28
? PRIOR APPLICATION NUMBER: PCT/US01/17800
? PRIOR FILING DATE: 2001-06-01
? PRIOR APPLICATION NUMBER: PCT/US01/19692
? PRIOR FILING DATE: 2001-06-20
? PRIOR APPLICATION NUMBER: PCT/US01/21066
? PRIOR FILING DATE: 2001-06-29
? PRIOR APPLICATION NUMBER: PCT/US01/21735
? PRIOR FILING DATE: 2001-07-09
? NUMBER OF SEQ ID NOS: 80
? SEQ ID NO 67
? LENGTH: 468
? TYPE: prt
? ORGANISM: Homo Sapien
US-10-036-342-67

```

Best Local Similarity: 32.6%; Pred. No. 88;
Matches 15; Conservative 8; Mismatches 9; Indels 14; Gaps 3;
Qy 6 KDDPKKEMEN-AG---NKITYMAA-----DKTISKLMY 37
Db 209 EEDPMKPTSWTENQAKIPEKVTYMAIQQDLAKGENDETVENTLT 254

RESULT 12
US-10-006-867-150
; Sequence 150, Application US/10006867
; Publication No. US20020119130A1
; GENERAL INFORMATION:
; APPLICANT: Baton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/006,867
; CURRENT FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: 60/063435
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/064215
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088740
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088811
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088825
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088863
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090688
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25

; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091628
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/096012
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: 60/096757
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 60/096949
; PRIOR FILING DATE: 1998-08-18
; PRIOR APPLICATION NUMBER: 60/096959
; PRIOR FILING DATE: 1998-08-18
; PRIOR APPLICATION NUMBER: 60/097954
; PRIOR FILING DATE: 1998-08-26
; PRIOR APPLICATION NUMBER: 60/097971
; PRIOR FILING DATE: 1998-08-26
; PRIOR APPLICATION NUMBER: 60/097979
; PRIOR FILING DATE: 1998-08-26
; PRIOR APPLICATION NUMBER: 60/098749
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/099741
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099763
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099792
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099812
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099815
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/100627
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100662
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100683
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/100684
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/100930
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/101279
; PRIOR FILING DATE: 1998-09-22
; PRIOR APPLICATION NUMBER: 60/101475
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101738
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101743
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101916
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/102570
; PRIOR FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: 60/103449
; PRIOR FILING DATE: 1998-10-06
; PRIOR APPLICATION NUMBER: 60/103678
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 60/103679
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 60/103711
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 60/105000
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: 60/105002
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: 60/105881
; PRIOR FILING DATE: 1998-10-27
; PRIOR APPLICATION NUMBER: 60/106030
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: 60/106464
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 60/106856
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: 60/108807

PRIOR FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: 60/112419
PRIOR FILING DATE: 1998-12-15
PRIOR APPLICATION NUMBER: 60/112422
PRIOR FILING DATE: 1998-12-15
PRIOR APPLICATION NUMBER: 60/112853
PRIOR FILING DATE: 1998-12-16
PRIOR APPLICATION NUMBER: 60/113011
PRIOR FILING DATE: 1998-12-16
PRIOR APPLICATION NUMBER: 60/112854
PRIOR FILING DATE: 1998-12-16
PRIOR APPLICATION NUMBER: 60/113300
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/113408
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/113430
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/113621
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/114223
PRIOR FILING DATE: 1998-12-30
PRIOR APPLICATION NUMBER: 60/115614
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/116527
PRIOR FILING DATE: 1999-01-20
PRIOR APPLICATION NUMBER: 60/116843
PRIOR FILING DATE: 1999-01-22
PRIOR APPLICATION NUMBER: 60/119285
PRIOR FILING DATE: 1999-02-09
PRIOR APPLICATION NUMBER: 60/119287
PRIOR FILING DATE: 1999-02-09
PRIOR APPLICATION NUMBER: 60/119525
PRIOR FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: 60/119549
PRIOR FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: 60/120014
PRIOR FILING DATE: 1999-02-11
PRIOR APPLICATION NUMBER: 60/129122
PRIOR FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: 60/129674
PRIOR FILING DATE: 1999-04-16
PRIOR APPLICATION NUMBER: 60/131291
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/138387
PRIOR FILING DATE: 1999-06-09
PRIOR APPLICATION NUMBER: 60/144791
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/169495
PRIOR FILING DATE: 1999-12-07
PRIOR APPLICATION NUMBER: 60/175481
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: 60/191007
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/199397
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/380139
PRIOR FILING DATE: 1998-08-25
PRIOR APPLICATION NUMBER: 09/311832
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 09/380137
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/380138
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/380142

Query Match 24.2% Score 52; DB 14; Length 468;
Best Local Similarity 32.6%; Pred. No. 88;
Matches 15; Conservative 8; Mismatches 9; Indels 14; Gaps 3;

Qy 6 KKDPSRKKEWMEN-AG---NKIYTMAA-----DKTISKLMT 37
Db 209 EBDPNKPTSWTENQAKIPEKVTPTMAAIQDGLAGKENDETVSNLT 254

RESULT 13
US-10-077-040-1
Sequence 1, Application US/10077040
Publication No. US20020156014A1
GENERAL INFORMATION:
APPLICANT: Lal, Preeti
Corley, Neil C.
Patterson, Chandra
TITLE OF INVENTION: HUMAN NEUROSECRETORY PROTEINS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/077,040
FILING DATE: 14-Feb-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/062,601
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Cerrone, Michael C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0510 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 468 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: ISLNOT01
CLONE: 2379427
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-077-040-1

Query Match 24.2% Score 52; DB 14; Length 468;
Best Local Similarity 32.6%; Pred. No. 88;
Matches 15; Conservative 8; Mismatches 9; Indels 14; Gaps 3;

Qy 6 KKDPSRKKEWMEN-AG---NKIYTMAA-----DKTISKLMT 37
Db 209 EBDPNKPTSWTENQAKIPEKVTPTMAAIQDGLAGKENDETVSNLT 254

RESULT 14
US-10-063-547-150
Sequence 150, Application US/10063547
Publication No. US20020182638A1
GENERAL INFORMATION:
APPLICANT: Baton, Dan L.
Filtavoff, Ellen
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.

```

: TITLE OF INVENTION: SECURED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE REFERENCE: P3230R1C1
: CURRENT APPLICATION NUMBER: US/10/063,547
: PRIOR FILING DATE: 2002-05-02
: PRIOR APPLICATION removed - See File Wrapper or Palm
: NUMBER OF SEQ ID NOS: 170
: SEQ ID NO 150
: LENGTH: 468
: TYPE: PRT
: ORGANISM: Homo Sapien
US-10-063-547-150

Query Match      24.2%  Score 52;  DB 14;  Length 468;
Best Local Similarity 32.6%  Pred. No. 88;
Matches 15;  Conservative 8;  Mismatches 9;  Indels 14;  Gaps 3

Qy      6  KKDSRKKEKMMEN-AG---NKITYMAA-----DKTISKMT 37
Db      209  EEDPKKPTSTWENQAGKIPEKVTYPMALIQDGLAKGNDVTSNTLT 254
          ::||::|||::|||::|||::|||::|||::|||::|||::|||

RESULT 15
US-10-036-041-67
: Sequence 67, Application US/10036041
: Publication No. US20020192751A1
: GENERAL INFORMATION:
: APPLICANT: Desnoyers, Luc
: APPLICANT: Ealton, Dan L.
: APPLICANT: Goddard, Audrey
: APPLICANT: Goddowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Pan, James
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Watanabe, Colin K.
: APPLICANT: Wood, William I.
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECURED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE REFERENCE: P3030R1C8
: CURRENT APPLICATION NUMBER: US/10/036,041
: CURRENT FILING DATE: 2001-12-26
: PRIOR APPLICATION NUMBER: 60/085579
: PRIOR FILING DATE: 1998-05-15
: PRIOR APPLICATION NUMBER: 60/112514
: PRIOR FILING DATE: 1998-12-15
: PRIOR APPLICATION NUMBER: 60/113300
: PRIOR FILING DATE: 1998-12-22
: PRIOR APPLICATION NUMBER: 60/114300
: PRIOR FILING DATE: 1998-12-23
: PRIOR APPLICATION NUMBER: 60/115605
: PRIOR FILING DATE: 1998-12-23
: PRIOR APPLICATION NUMBER: 60/11621
: PRIOR FILING DATE: 1998-12-23
: PRIOR APPLICATION NUMBER: 60/114140
: PRIOR FILING DATE: 1998-12-23
: PRIOR APPLICATION NUMBER: 60/115552
: PRIOR FILING DATE: 1999-01-12
: PRIOR APPLICATION NUMBER: 60/116843
: PRIOR FILING DATE: 1999-01-22
: PRIOR APPLICATION NUMBER: 60/125774
: PRIOR FILING DATE: 1999-03-23
: PRIOR APPLICATION NUMBER: 60/125778
: PRIOR FILING DATE: 1999-03-23
: PRIOR APPLICATION NUMBER: 60/125826
: PRIOR FILING DATE: 1999-03-24
: PRIOR APPLICATION NUMBER: 60/127035
: PRIOR FILING DATE: 1999-03-31
: PRIOR APPLICATION NUMBER: 60/127706
: PRIOR FILING DATE: 1999-04-05
: PRIOR APPLICATION NUMBER: 60/129122
: PRIOR FILING DATE: 1999-04-13
: PRIOR APPLICATION NUMBER: 60/130359

```

PRIOR	FILING DATE:	1999-04-21
PRIOR	APPLICATION NUMBER:	60/131270
PRIOR	FILING DATE:	1999-04-22
PRIOR	APPLICATION NUMBER:	60/131272
PRIOR	FILING DATE:	1999-04-27
PRIOR	APPLICATION NUMBER:	60/131291
PRIOR	FILING DATE:	1999-04-27
PRIOR	APPLICATION NUMBER:	60/132371
PRIOR	FILING DATE:	1999-05-04
PRIOR	APPLICATION NUMBER:	60/132379
PRIOR	FILING DATE:	1999-05-04
PRIOR	APPLICATION NUMBER:	60/132383
PRIOR	FILING DATE:	1999-05-04
PRIOR	APPLICATION NUMBER:	60/135750
PRIOR	FILING DATE:	1999-05-25
PRIOR	APPLICATION NUMBER:	60/138166
PRIOR	FILING DATE:	1999-06-08
PRIOR	APPLICATION NUMBER:	60/144191
PRIOR	FILING DATE:	1999-07-20
PRIOR	APPLICATION NUMBER:	60/146970
PRIOR	FILING DATE:	1999-08-03
PRIOR	APPLICATION NUMBER:	60/162506
PRIOR	FILING DATE:	1999-10-29
PRIOR	APPLICATION NUMBER:	09/311832
PRIOR	FILING DATE:	1999-05-14
PRIOR	APPLICATION NUMBER:	09/380142
PRIOR	FILING DATE:	1999-08-25
PRIOR	APPLICATION NUMBER:	09/644848
PRIOR	FILING DATE:	2000-08-22
PRIOR	APPLICATION NUMBER:	09/747259
PRIOR	FILING DATE:	2000-12-20
PRIOR	APPLICATION NUMBER:	09/816744
PRIOR	FILING DATE:	2001-03-22
PRIOR	APPLICATION NUMBER:	09/854208
PRIOR	FILING DATE:	2001-05-10
PRIOR	APPLICATION NUMBER:	09/854280
PRIOR	FILING DATE:	2001-05-10
PRIOR	APPLICATION NUMBER:	09/874503
PRIOR	FILING DATE:	2001-06-05
PRIOR	APPLICATION NUMBER:	09/869599
PRIOR	FILING DATE:	2001-06-29
PRIOR	APPLICATION NUMBER:	09/908,827
PRIOR	FILING DATE:	2001-07-18
PRIOR	APPLICATION NUMBER:	PCT/US99/107333
PRIOR	FILING DATE:	1999-05-14
PRIOR	APPLICATION NUMBER:	PCT/US99/285531
PRIOR	FILING DATE:	1999-12-02
PRIOR	APPLICATION NUMBER:	PCT/US99/307200
PRIOR	FILING DATE:	1999-12-22
PRIOR	APPLICATION NUMBER:	PCT/US00/056010
PRIOR	FILING DATE:	2000-03-01
PRIOR	APPLICATION NUMBER:	PCT/US00/058411
PRIOR	FILING DATE:	2000-03-02
PRIOR	APPLICATION NUMBER:	PCT/US00/140422
PRIOR	FILING DATE:	2000-05-22
PRIOR	APPLICATION NUMBER:	PCT/US00/152660
PRIOR	FILING DATE:	2000-06-02
PRIOR	APPLICATION NUMBER:	PCT/US00/235222
PRIOR	FILING DATE:	2000-08-23
PRIOR	APPLICATION NUMBER:	PCT/US00/233282
PRIOR	FILING DATE:	2000-08-24
PRIOR	APPLICATION NUMBER:	PCT/US00/326787
PRIOR	FILING DATE:	2000-12-01
PRIOR	APPLICATION NUMBER:	PCT/US00/349565
PRIOR	FILING DATE:	2000-12-20
PRIOR	APPLICATION NUMBER:	PCT/US01/065200
PRIOR	FILING DATE:	2001-02-28
PRIOR	APPLICATION NUMBER:	PCT/US01/178000
PRIOR	FILING DATE:	2001-06-01
PRIOR	APPLICATION NUMBER:	PCT/US01/196922
PRIOR	FILING DATE:	2001-06-20
PRIOR	APPLICATION NUMBER:	PCT/US01/210666
PRIOR	FILING DATE:	2001-06-29

; PRIOR APPLICATION NUMBER: PCT/US01/21735
 ; PRIOR FILING DATE: 2001-07-09
 ; NUMBER OF SEQ ID NOS: 80
 ; SEQ ID NO 67
 ; LENGTH: 468
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 US-10-036-041-67

Query Match 24.2%; Score 52; DB 14; Length 468;
 Best Local Similarity 32.6%; Pred. No. 88;
 Matches 15; Conservative 8; Mismatches 9; Indels 14; Gaps 3;

Oy 6 KDPSPKKEWMEN-AG---NKIYTWAA-----DKTISKLMY 37
 ::||::| ||| |
 Db 209 EEDPVKPTSWTENQAGKIPEKVTPMALIQDGLAKGENDETYSNTLT 254

Search completed: July 24, 2003, 12:18:20
 Job time : 6.82665 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 11:49:53 ; Search time 60.2635 Seconds
(without alignments)
577.701 Million cell updates/sec

Title: US-09-991-681-30

Perfect score: 215
Sequence: 1 SPKVEKDPKRKKEMENAGNKIYTMADTKTISLMTYEK 40

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 5580241 seqs, 870357830 residues

Total number of hits satisfying chosen parameters: 5580241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA_Main:*

1:	/cgn2_6/ptodata/1/paa/PCRTUS_COMB.pep.*
2:	/cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3:	/cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4:	/cgn2_6/ptodata/1/paa/US08_COMB.pep.*
5:	/cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6:	/cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7:	/cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8:	/cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9:	/cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10:	/cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11:	/cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12:	/cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13:	/cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14:	/cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15:	/cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16:	/cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17:	/cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18:	/cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19:	/cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20:	/cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21:	/cgn2_6/ptodata/1/paa/US097A_COMB.pep.*
22:	/cgn2_6/ptodata/1/paa/US097B_COMB.pep.*
23:	/cgn2_6/ptodata/1/paa/US098_COMB.pep.*
24:	/cgn2_6/ptodata/1/paa/US099_COMB.pep.*
25:	/cgn2_6/ptodata/1/paa/US099B_COMB.pep.*
26:	/cgn2_6/ptodata/1/paa/US100_COMB.pep.*
27:	/cgn2_6/ptodata/1/paa/US101_COMB.pep.*
28:	/cgn2_6/ptodata/1/paa/US102_COMB.pep.*
29:	/cgn2_6/ptodata/1/paa/US103_COMB.pep.*
30:	/cgn2_6/ptodata/1/paa/US104_COMB.pep.*
31:	/cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	215	100.0	40	12 US-08-842-385-9	Sequence 9, Appl1
2	215	100.0	40	25 US-09-991-681-30	Sequence 30, Appl1

3	215	100.0	141	28 US-10-221-279-7783	Sequence 7783, Ap
4	215	100.0	467	12 US-08-842-385-6	Sequence 6, Appl1
5	215	100.0	518	25 US-09-991-681-27	Sequence 27, Appl1
6	215	100.0	1770	1 PCT-US03-01943-44	Sequence 44, Appl1
7	215	100.0	1770	US-10-144-198-44	Sequence 44, Appl1
8	215	100.0	1807	1 PCT-US01-08631-40087	Sequence 40087, A
9	215	100.0	1839	1 PCT-US01-42950-495	Sequence 495, App
10	215	100.0	1839	30 US-10-416-993-495	Sequence 495, App
11	215	100.0	1872	1 PCT-US03-04508-32	Sequence 32, Appl1
12	215	100.0	1982	1 PCT-US01-08631-40090	Sequence 40090, A
13	215	100.0	2221	1 PCT-US03-01943-30	Sequence 30, Appl1
14	215	100.0	2221	27 US-10-144-198-30	Sequence 30, Appl1
15	65	30.2	1065	30 US-10-437-963-153423	Sequence 153423, A
16	64	29.8	621	22 US-09-791-537-91859	Sequence 22859, A
17	63.5	29.5	604	31 US-60-360-039-22757	Sequence 22757, A
18	59.5	27.7	232	13 US-08-993-002A-5796	Sequence 5796, Ap
19	59.5	27.7	232	10 US-08-625-811-1380	Sequence 1380, Ap
20	59.5	27.7	233	13 US-08-993-002A-5795	Sequence 5797, Ap
21	59.5	27.7	388	1 PCT-US02-03987-11623	Sequence 11623, A
22	59.5	27.7	388	13 US-08-993-002A-5797	Sequence 5797, Ap
23	59.5	27.7	388	23 US-09-815-242-11623	Sequence 11623, A
24	59.5	27.7	388	26 US-10-072-851-11623	Sequence 11623, A
25	59	27.4	264	31 US-60-360-039-11489	Sequence 11489, A
26	59	27.4	264	31 US-60-360-039-14338	Sequence 14338, A
27	59	27.4	264	31 US-60-360-039-14597	Sequence 14597, A
28	59	27.4	264	31 US-60-360-039-14979	Sequence 14979, A
29	59	27.4	277	21 US-09-739-449-9886	Sequence 9886, Ap
30	59	27.4	277	23 US-09-803-110-9886	Sequence 9886, Ap
31	59	27.4	314	28 US-10-219-999-37532	Sequence 37532, A
32	59	27.4	314	30 US-10-425-114-44366	Sequence 44366, A
33	59	27.4	314	31 US-60-324-109-24021	Sequence 24021, A
34	57	26.5	275	30 US-10-431-652-5591	Sequence 5591, Ap
35	57	26.5	664	27 US-10-181-612-2	Sequence 2, Appl1
36	56.5	26.3	158	26 US-10-053-853A-1179	Sequence 1179, Ap
37	56.5	26.3	388	1 PCT-US02-03987-11458	Sequence 11458, A
38	56.5	26.3	388	12 PCT-US98-06371-350	Sequence 350, App
39	56.5	26.3	388	12 US-08-833-457-350	Sequence 350, App
40	56.5	26.3	388	23 US-09-815-242-11458	Sequence 11458, A
41	56.5	26.3	388	23 US-09-815-242-11458	Sequence 350, App
42	56.5	26.3	388	26 US-10-072-851-11458	Sequence 11458, A
43	56.5	26.3	388	28 US-10-282-122A-55031	Sequence 55031, A
44	56	26.0	219	19 US-09-513-996A-55105	Sequence 55105, A
45	56	26.0	219	19 US-09-513-996A-63822	Sequence 63822, A

ALIGNMENTS

RESULT 1
US-08-842-385-9
Sequence 9, Application US/08842385
GENERAL INFORMATION:
APPLICANT: Russell, John
APPLICANT: Colpitts, Tracey
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
FOR DETECTING DISEASE OF THE PROSTATE
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESSES:
ADDRESS: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/842,385
FILING DATE:
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Potembaki, Priscilla E
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 6084.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/937-6365
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: None
US-08-842-385-9

Query Match 100.0%; Score 215; DB 12; Length 40;
Best Local Similarity 100.0%; Pred. No. 1.5e-20;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SPKVEKKDPSRKKEWENAGNKIYTMADKTIISKLMTEYK 40
DB 1 SPKVEKKDPSRKKEWENAGNKIYTMADKTIISKLMTEYK 40

RESULT 2
US-09-991-681-30
Sequence 30, Application US/09991681
GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA
COHEN, MAURICE
COLPITTS, TRACEY L.
FRIEDMAN, PAULA N.
GORDON, JULIAN
GRANADOS, EDWARD N.
HODGES, STEVEN C.
KLASS, MICHAEL R.
KRATOCHVIL, JON D.
ROBERTS-RAPP, LISA
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESS: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/991,681
FILING DATE: 26-Nov-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/065,383
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6084.US.PI
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: None
SEQUENCE DESCRIPTION: SEQ ID NO: 30:
US-09-991-681-30

Query Match 100.0%; Score 215; DB 25; Length 40;
Best Local Similarity 100.0%; Pred. No. 1.5e-20;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SPKVEKKDPSRKKEWENAGNKIYTMADKTIISKLMTEYK 40
DB 1 SPKVEKKDPSRKKEWENAGNKIYTMADKTIISKLMTEYK 40

RESULT 3
US-10-221-279-7783
Sequence 7783, Application US/10221279
GENERAL INFORMATION:
APPLICANT: Hysed, Inc
TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
FILE REFERENCE: 21272-046
CURRENT APPLICATION NUMBER: US/10/221,279
PRIOR FILING DATE: 2002-09-06
PRIOR APPLICATION NUMBER: 09/574,454
PRIOR FILING DATE: 2000-05-19
PRIOR APPLICATION NUMBER: 09/519,705
PRIOR FILING DATE: 2000-03-07
NUMBER OF SEQ ID NOS: 12360
SOFTWARE: Custom
SEQ ID NO 7783
LENGTH: 141
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(141)
OTHER INFORMATION: Xaa = any amino acid or nothing
US-10-221-279-7783

Query Match 100.0%; Score 215; DB 28; Length 141;
Best Local Similarity 100.0%; Pred. No. 6.8e-20;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SPKVEKKDPSRKKEWENAGNKIYTMADKTIISKLMTEYK 40
DB 96 SPKVEKKDPSRKKEWENAGNKIYTMADKTIISKLMTEYK 135

RESULT 4
US-08-842-385-6
Sequence 6, Application US/08842385
GENERAL INFORMATION:
APPLICANT: Russell, John
APPLICANT: Colpitts, Tracey
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
FOR DETECTING DISEASE OF THE PROSTATE
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESS: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/842,385
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Porembski, Priscilla E
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 6084.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/937-6365
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: None
US-08-842-385-6

Query Match 100.0%; Score 215; DB 12; Length 467;
Best Local Similarity 100.0%; Pred. No. 2,9e-19;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPKVEKDPSPRKEMWENAGNKIYTMADTKISKLMTYK 40
Db 281 SPKVEKDPSPRKEMWENAGNKIYTMADTKISKLMTYK 320

RESULT 5
US-09-991-681-27
Sequence 27, Application US/09991681
GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA
COHEN, MAURICE
COLPITTS, TRACEY L.
FRIEDMAN, PAULA N.
GORDON, JULIAN
GRANADOS, EDWARD N.
HODGES, STEVEN C.
KLASS, MICHAEL R.
KRATOCHVIL, JON D.
ROBERTS-RAPP, LISA
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/991,681
FILING DATE: 26-Nov-2001
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/065,383
FILING DATE: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441

REFERENCE/DOCKET NUMBER: 6084.US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX: <unknown>
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 518 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: None
SEQUENCE DESCRIPTION: SEQ ID NO: 27:
US-09-991-681-27

Query Match 100.0%; Score 215; DB 25; Length 518;
Best Local Similarity 100.0%; Pred. No. 3.3e-19;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPKVEKDPSPRKEMWENAGNKIYTMADTKISKLMTYK 40
Db 332 SPKVEKDPSPRKEMWENAGNKIYTMADTKISKLMTYK 371

RESULT 6
PCT-US03-01943-44
Sequence 44, Application PC/TUS0301943
GENERAL INFORMATION:
APPLICANT: ORIGENE TECHNOLOGIES INC
TITLE OF INVENTION: CANCER GENES
FILE REFERENCE: 3U 9U 901 PCT
CURRENT APPLICATION NUMBER: PCT/US03/01943
CURRENT FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: US 10/054,935
PRIOR FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: US 60/356,130
PRIOR FILING DATE: 2002-02-14
PRIOR APPLICATION NUMBER: US 10/102,946
PRIOR FILING DATE: 2002-03-22
PRIOR APPLICATION NUMBER: US 10/117,229
PRIOR FILING DATE: 2002-04-08
PRIOR APPLICATION NUMBER: US 10/144,198
PRIOR FILING DATE: 2002-05-14
PRIOR APPLICATION NUMBER: US 10/197,824
PRIOR FILING DATE: 2002-07-19
NUMBER OF SEQ ID NOS: 102
SOFTWARE: PatentIn version 3.1
SEQ ID NO 44
LENGTH: 1770
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US03-01943-44

Query Match 100.0%; Score 215; DB 1; Length 1770;
Best Local Similarity 100.0%; Pred. No. 1.4e-18;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPKVEKDPSPRKEMWENAGNKIYTMADTKISKLMTYK 40
Db 1584 SPKVEKDPSPRKEMWENAGNKIYTMADTKISKLMTYK 1623

RESULT 7
US-10-144-198-44
Sequence 44, Application US/10144198
GENERAL INFORMATION:
APPLICANT: Origene Technologies Inc
TITLE OF INVENTION: Regulated Prostate Cance Genes
FILE REFERENCE: 9U 105 R1
CURRENT APPLICATION NUMBER: US/10/144,198
CURRENT FILING DATE: 2002-05-14
NUMBER OF SEQ ID NOS: 44
SOFTWARE: PatentIn version 3.0

SEQ ID NO 44
LENGTH: 1770
TYPE: PRT
ORGANISM: Homo sapiens
US-10-144-198-44

Query Match 100.0%; Score 215; DB 27; Length 1770;
Best Local Similarity 100.0%; Pred. No. 1.4e-18;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPKVEKKDPSRKKEWENAGNKIYTMAADKTIISKLMTEYK 40
Db 1584 SPKVEKKDPSRKKEWENAGNKIYTMAADKTIISKLMTEYK 1623

RESULT 8
PCT-US01-08631-40087
Sequence 40087, Application PC/TUS0108631
GENERAL INFORMATION:

APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 21272-049
CURRENT APPLICATION NUMBER: PCT/US01/08631
CURRENT FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 09/540,217
PRIOR FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: 09/649,167
PRIOR FILING DATE: 2000-08-23
NUMBER OF SEQ ID NOS: 60736
SOFTWARE: Custom
SEQ ID NO 40087
LENGTH: 1807
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (48)..(62)
OTHER INFORMATION: Fusion glycoprotein F0 domain identified by eMATRIX,
OTHER INFORMATION: accession number PF00523D, P-value=7.188e-10, raw score of 11.39
NAME/KEY: DOMAIN
LOCATION: (941)..(950)
OTHER INFORMATION: Helper component proteinase domain identified by Pfam,
OTHER INFORMATION: accession name Pepsidase_C6, E-value=0.0056, Pfam score of 7.9
PCT-US01-08631-40087

Query Match 100.0%; Score 215; DB 1; Length 1807;
Best Local Similarity 100.0%; Pred. No. 1.5e-18;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPKVEKKDPSRKKEWENAGNKIYTMAADKTIISKLMTEYK 40
Db 1621 SPKVEKKDPSRKKEWENAGNKIYTMAADKTIISKLMTEYK 1660

RESULT 9
PCT-US01-42950-495
Sequence 495, Application PC/TUS0142950
GENERAL INFORMATION:

APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 21272-096
CURRENT APPLICATION NUMBER: PCT/US01/42950
CURRENT FILING DATE: 2001-11-16
PRIOR APPLICATION NUMBER: 09/714,936
PRIOR FILING DATE: 2000-11-17
NUMBER OF SEQ ID NOS: 682
SOFTWARE: PatentIn version 3.0
SEQ ID NO 495
LENGTH: 1839
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US01-42950-495

Query Match 100.0%; Score 215; DB 1; Length 1839;
Best Local Similarity 100.0%; Pred. No. 1.5e-18;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPKVEKKDPSRKKEWENAGNKIYTMAADKTIISKLMTEYK 40
Db 1653 SPKVEKKDPSRKKEWENAGNKIYTMAADKTIISKLMTEYK 1692

RESULT 10
US-10-416-993-495
Sequence 495, Application US/10416993
GENERAL INFORMATION:

APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 21272-096
CURRENT APPLICATION NUMBER: US/10/416,993
CURRENT FILING DATE: 2003-11-16
PRIOR APPLICATION NUMBER: 09/714,936
PRIOR FILING DATE: 2000-11-17
NUMBER OF SEQ ID NOS: 682
SOFTWARE: PatentIn version 3.0
SEQ ID NO 495
LENGTH: 1839
TYPE: PRT
ORGANISM: Homo sapiens
US-10-416-993-495

Query Match 100.0%; Score 215; DB 30; Length 1839;
Best Local Similarity 100.0%; Pred. No. 1.5e-18;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPKVEKKDPSRKKEWENAGNKIYTMAADKTIISKLMTEYK 40
Db 1653 SPKVEKKDPSRKKEWENAGNKIYTMAADKTIISKLMTEYK 1692

RESULT 11
PCT-US03-04508-32
Sequence 32, Application PC/TUS0304508
GENERAL INFORMATION:

APPLICANT: IDEC PHARMACEUTICALS
TITLE OF INVENTION: GATELY, DENNIS
TITLE OF INVENTION: PROSTATE SPECIFIC GENES AND THE USE THEREOF IN DESIGN
TITLE OF INVENTION: OR THERAPEUTICS
FILE REFERENCE: 037003/0301985
CURRENT APPLICATION NUMBER: PCT/US03/04508
CURRENT FILING DATE: 2003-02-19
PRIOR APPLICATION NUMBER: 60/357,140
PRIOR FILING DATE: 2002-02-19
PRIOR APPLICATION NUMBER: 60/396,082
PRIOR FILING DATE: 2002-07-17
PRIOR APPLICATION NUMBER: 60/386,759
PRIOR FILING DATE: 2002-06-10
NUMBER OF SEQ ID NOS: 93
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 32
LENGTH: 1872
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US03-04508-32

Query Match 100.0%; Score 215; DB 1; Length 1872;
Best Local Similarity 100.0%; Pred. No. 1.5e-18;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPKVEKKDPSRKKEWENAGNKIYTMAADKTIISKLMTEYK 40
Db 1686 SPKVEKKDPSRKKEWENAGNKIYTMAADKTIISKLMTEYK 1725

RESULT 12
PCT-US01-08631-40090

```
; Sequence 40090, Application PC/TUS0108631
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-049
; CURRENT APPLICATION NUMBER: PCT/US01/08631
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 40090
; LENGTH: 1982
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (11)..(25)
; OTHER INFORMATION: Fusion glycoprotein F0 domain identified by eMATRIX.
; OTHER INFORMATION: accession number PF00523D, p-value=7.188e-10, raw score of 11.39
; NAME/KEY: DOMAIN
; LOCATION: (1065)..(1074)
; OTHER INFORMATION: Helper component proteinase domain identified by Pfam,
; OTHER INFORMATION: accession name Peptidase_C6, E-value=0.0056, Pfam score of 7.9
PCT-US01-08631-40090
```

```
Query Match 100.0%; Score 215; DB 1; Length 1982;
Best Local Similarity 100.0%; Pred. No. 1.7e-18;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 SPKVEKKDPSRKKEWENAGNKIYTMADKTIISKLTMEYK 40
Db 1735 SPKVEKKDPSRKKEWENAGNKIYTMADKTIISKLTMEYK 1774
```

```
RESULT 13
PCT-US03-01943-30
; Sequence 30, Application PC/TUS0301943
; GENERAL INFORMATION:
; APPLICANT: ORIGENE TECHNOLOGIES INC
; TITLE OF INVENTION: CANCER GENES
; FILE REFERENCE: 3U 9U 901 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/01943
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: US 10/054,935
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: US 60/356,130
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 10/102,946
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: US 10/117,229
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: US 10/144,198
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: US 10/197,824
; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30
; LENGTH: 2221
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-01943-30
```

```
Query Match 100.0%; Score 215; DB 1; Length 2221;
Best Local Similarity 100.0%; Pred. No. 1.9e-18;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 SPKVEKKDPSRKKEWENAGNKIYTMADKTIISKLTMEYK 40
Db 2035 SPKVEKKDPSRKKEWENAGNKIYTMADKTIISKLTMEYK 2074
```

```
RESULT 14
US-10-144-198-30
; Sequence 30, Application US/10144198
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies Inc
; TITLE OF INVENTION: Regulated Prostate Cance Genes
; FILE REFERENCE: 9U 105 R1
; CURRENT APPLICATION NUMBER: US/10/144,198
; CURRENT FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 30
; LENGTH: 2221
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-198-30
```

```
Query Match 100.0%; Score 215; DB 27; Length 2221;
Best Local Similarity 100.0%; Pred. No. 1.9e-18;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 SPKVEKKDPSRKKEWENAGNKIYTMADKTIISKLTMEYK 40
Db 2035 SPKVEKKDPSRKKEWENAGNKIYTMADKTIISKLTMEYK 2074
```

```
RESULT 15
US-10-437-963-153423
; Sequence 153423, Application US/10437963
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 153423
; LENGTH: 1065
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1065)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_53380C.1.pcp
US-10-437-963-153423
```

```
Query Match 30.2%; Score 65; DB 30; Length 1065;
Best Local Similarity 32.3%; Pred. No. 55;
Matches 10; Conservative 8; Mismatches 13; Indels 0; Gaps 0;
```

```
Qy 1 SPKVEKKDPSRKKEWENAGNKIYTMADKT 31
Db 488 TPLSDSDLRCCOMNNNAKRVATLEPDR 518
```

```
Search completed: July 24, 2003, 12:15:14
Job time : 61.2635 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: July 24, 2003, 11:54:33 ; Search time 1.1713 Seconds
(without alignments)
140.227 Million cell updates/sec

Title: US-09-991-681-30

Perfect score: 215
Sequence: 1 SPKYEKKDPSRKKEWMENAGNKIYTMADKTISKLMTEYK 40

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 41799 seqs, 4106219 residues

Total number of hits satisfying chosen parameters: 41799

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB pep:*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB pep:*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB pep:*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB pep:*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB pep:*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB pep:*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	215	100.0	1872	US-10-367-978-32	Sequence 32, Appl
2	53	24.7	489	US-10-294-433-250	Sequence 250, App
3	52	24.2	477	US-10-273-573-7140	Sequence 7140, Ap
4	51	23.7	4544	US-10-464-368-68	Sequence 68, Appl
5	51	23.7	4545	US-10-464-368-67	Sequence 67, Appl
6	51	23.7	4545	US-10-464-368-71	Sequence 71, Appl
7	50	23.3	1469	US-60-479-073-335	Sequence 335, App
8	47	21.9	98	US-10-029-988B-42	Sequence 42, Appl
9	47	21.9	98	US-10-032-037B-42	Sequence 42, Appl
10	46	21.4	329	US-09-290-586A-19	Sequence 19, Appl
11	46	21.4	1101	US-10-287-971-18	Sequence 18, Appl
12	45.5	21.2	1839	US-10-273-573-9476	Sequence 9476, Ap
13	44.5	20.7	606	US-10-331-946A-60	Sequence 60, Appl
14	44	20.5	98	US-10-029-988B-41	Sequence 41, Appl
15	44	20.5	98	US-10-032-037B-41	Sequence 41, Appl
16	44	20.5	98	US-10-308-817-42	Sequence 42, Appl
17	44	20.5	500	US-10-014-099F-57	Sequence 57, Appl
18	44	20.5	854	US-10-273-573-6450	Sequence 6450, Ap
19	43.5	20.2	462	US-10-273-573-7554	Sequence 7554, Ap
20	43.5	20.2	1152	US-60-479-073-81	Sequence 81, Appl
21	43.5	20.2	1118	US-60-479-073-79	Sequence 79, Appl
22	43.5	20.2	1231	US-60-479-073-77	Sequence 77, Appl
23	43	20.0	40	US-10-273-573-10766	Sequence 10766, A
24	43	20.0	524	US-60-478-186-3402	Sequence 3202, Ap
25	43	20.0	939	PCT-US03-09929-32	Sequence 32, Appl
26	42.5	19.8	508	US-10-465-302-13	Sequence 13, Appl

27	42.5	19.8	950	US-10-273-573-10847	Sequence 10847, A
28	42	19.5	18	PCT-US03-18896-353	Sequence 353, App
29	42	19.5	18	US-10-462-262-353	Sequence 353, App
30	42	19.5	30	PCT-US03-18896-381	Sequence 381, App
31	42	19.5	30	US-10-462-262-381	Sequence 381, App
32	42	19.5	331	US-09-290-586A-20	Sequence 20, Appl
33	42	19.5	506	PCT-US03-20480-31	Sequence 31, Appl
34	42	19.5	583	US-10-433-802-13	Sequence 13, Appl
35	42	19.5	613	US-10-014-099F-21	Sequence 21, Appl
36	42	19.5	620	US-10-014-099F-23	Sequence 23, Appl
37	42	19.5	1227	US-09-291-417D-105	Sequence 105, App
38	42	19.5	4660	US-10-464-368-74	Sequence 74, Appl
39	42	19.5	5065	PCT-US02-18638A-52	Sequence 52, Appl
40	41.5	19.3	222	US-10-273-573-10263	Sequence 10263, A
41	41.5	19.3	548	US-10-450-200-2	Sequence 2, Appl
42	41	19.1	123	US-10-273-573-9406	Sequence 9406, Ap
43	41	19.1	173	US-10-351-161A-4	Sequence 4, Appl
44	41	19.1	242	US-10-273-573-8424	Sequence 8424, Ap
45	41	19.1	267	US-10-273-573-9399	Sequence 9399, Ap

ALIGNMENTS

```
RESULT 1
US-10-367-978-32
; Sequence 32, Application US/10367978
; GENERAL INFORMATION:
; APPLICANT: GATELY, DENNIS
; TITLE OF INVENTION: PROSTATE SPECIFIC GENES AND THE USE THEREOF IN DESIGN
; FILE REFERENCE: 037003-0301988
; CURRENT APPLICATION NUMBER: US/10/367, 978
; CURRENT FILING DATE: 2003-02-19
; PRIOR APPLICATION NUMBER: 60/357,140
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: 60/396,082
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 60/386,759
; PRIOR FILING DATE: 2002-06-10
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 1872
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-367-978-32

Query Match      100.0%; Score 215; DB 6; Length 1872;
Best Local Similarity 100.0%; Pred. No. 3.3e-19;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1 SPKYEKKDPSRKKEWMENAGNKIYTMADKTISKLMTEYK 40
Db      1686 SPKYEKKDPSRKKEWMENAGNKIYTMADKTISKLMTEYK 1725

RESULT 2
US-10-294-433-250
; Sequence 250, Application US/10294433
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 792CIP4
; CURRENT APPLICATION NUMBER: US/10/294, 433
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: PCT/US01/14826
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/989, 600
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 09/577, 408
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: 10/115, 831
```

```

; PRIOR FILING DATE: 2002-04-02
; PRIOR APPLICATION NUMBER: 09/677,298
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/695,781
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 10/150,802
; PRIOR FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: 09/715,869
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 10/167,379
; PRIOR FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 09/775,330
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 864
; SOFTWARE: Custom
; SEQ ID NO 250
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-294-433-250
```

```

Query Match          24.7%; Score 53; DB 6; Length 489;
Best Local Similarity 36.4%; Pred. No. 9.3;
Matches 16; Conservative 5; Mismatches 15; Indels 8; Gaps 3;
```

```

QY      3 KYEKKDPSRKKE---W---WENAGNKIYTMADKTI SKLMTYK 40
DB      345 KMTKKIKKEKTIITWRTWENNKKALLQWAEKTVRD--KEKK 386
```

```

RESULT 3
US-10-273-573-7140
; Sequence 7140, Application US/10273573
; GENERAL INFORMATION:
; APPLICANT: HySeq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: US/10/273,573
; PRIOR FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 7140
; LENGTH: 477
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: DOMAIN
; LOCATION: (34)..(52)
; OTHER INFORMATION: GLUCOSE/RIBITOL DEHYDROGENASE FAMILY SIGNATURE domain
; OTHER INFORMATION: identified by eMATRIX, accession number PR00081A, p-value=6.22e-
; OTHER INFORMATION: 13, raw score of 10.53
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (33)..(218)
; OTHER INFORMATION: short chain dehydrogenase domain identified by Pfam,
; OTHER INFORMATION: accession name adh_short, E-value=1.5e-60, Pfam score of 214.6
US-10-273-573-7140
```

```

Query Match          24.2%; Score 52; DB 6; Length 477;
Best Local Similarity 25.6%; Pred. No. 12;
Matches 10; Conservative 11; Mismatches 18; Indels 0; Gaps 0;
```

```

QY      2 PYEKKDPSRKKEWENAGNKIYTMADKTI SKLMTYK 40
DB      386 PKYKSSPAREQWENKEDDLTVSFRSVLTNTYQK 424
```

```

RESULT 4
US-10-464-368-68
```

```

; Sequence 68, Application US/10464368
; GENERAL INFORMATION:
; APPLICANT: Krumlauf, Robb
; APPLICANT: Ellices, Debra
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REGULATING BONE DEPOSITION
; FILE REFERENCE: 40716-1P-017
; CURRENT APPLICATION NUMBER: US/10/464,368
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: 60/388,970
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 68
; LENGTH: 4544
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-10-464-368-68
```

```

Query Match          23.7%; Score 51; DB 6; Length 4544;
Best Local Similarity 36.4%; Pred. No. 1.6e+02;
Matches 12; Conservative 6; Mismatches 13; Indels 2; Gaps 1;
```

```

QY      8 DPSRKKEWENAGN--KIYTMADKTI SKLMTYK 38
DB      4009 DPLRGTWYSDWGNHPKIEFTAMDGTLRETLVQ 4041
```

```

RESULT 5
US-10-464-368-67
; Sequence 67, Application US/10464368
; GENERAL INFORMATION:
; APPLICANT: Krumlauf, Robb
; APPLICANT: Ellices, Debra
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REGULATING BONE DEPOSITION
; FILE REFERENCE: 40716-1P-017
; CURRENT APPLICATION NUMBER: US/10/464,368
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: 60/388,970
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 67
; LENGTH: 4545
; TYPE: PRT
; ORGANISM: MOUSE
US-10-464-368-67
```

```

Query Match          23.7%; Score 51; DB 6; Length 4545;
Best Local Similarity 36.4%; Pred. No. 1.6e+02;
Matches 12; Conservative 6; Mismatches 13; Indels 2; Gaps 1;
```

```

QY      8 DPSRKKEWENAGN--KIYTMADKTI SKLMTYK 38
DB      4010 DPLRGTWYSDWGNHPKIEFTAMDGTLRETLVQ 4042
```

```

RESULT 6
US-10-464-368-71
; Sequence 71, Application US/10464368
; GENERAL INFORMATION:
; APPLICANT: Krumlauf, Robb
; APPLICANT: Ellices, Debra
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REGULATING BONE DEPOSITION
; FILE REFERENCE: 40716-1P-017
; CURRENT APPLICATION NUMBER: US/10/464,368
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: 60/388,970
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 71
; LENGTH: 4545
; TYPE: PRT
```



```
; ORGANISM: MOUSE
US-10-464-368-71

Query Match      23.7% Score 51; DB 6; Length 4545;
Best Local Similarity 36.4%; Pred. No. 1.6e+02;
Matches 12; Conservative 6; Mismatches 13; Indels 2; Gaps 1;

OY      8 DPSRKKEWMENAGN--KIYTMADKTIKSLMTE 38
DB      4010 DPLRGTYWSDWGNHPKTIETAMDGTLRRTLVQ 4042

RESULT 7
US-60-479-073-335
; Sequence 335, Application US/60479073
; GENERAL INFORMATION:
; APPLICANT: De Wilde, Gert Jules Hector
; APPLICANT: Saunders, Michael John Scott
; APPLICANT: Logghe, Marc Georges
; TITLE OF INVENTION: Amino acid sequences useful for developing compounds for the
; TITLE OF INVENTION: prevention and/or treatment of metabolic diseases and nucleotide
; TITLE OF INVENTION: sequences encoding such amino acid sequences.
; FILE REFERENCE: D00590.70042.US
; CURRENT APPLICATION NUMBER: US/60/479,073
; CURRENT FILING DATE: 2003-06-17
; NUMBER OF SEQ ID NOS: 526
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 335
; LENGTH: 1469
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-60-479-073-335

Query Match      23.3% Score 50; DB 7; Length 1469;
Best Local Similarity 44.8%; Pred. No. 68;
Matches 13; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

OY      4 VEKDPSSRKKEWMENAGNKIYTMADKTI 32
DB      365 LEAKDEVKKKETHERSLNSIVTELNKR 393

RESULT 8
US-10-029-988B-42
; Sequence 42, Application US/10029988B
; GENERAL INFORMATION:
; APPLICANT: Bio-Technology General Corp.
; TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
; TITLE OF INVENTION: MOIETIES, ANTIBODIES TO SUCH EPITOPES, AND USES THEREOF
; FILE REFERENCE: 10793/46
; CURRENT APPLICATION NUMBER: US/10/029,988B
; CURRENT FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/258,948
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 204
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 42
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-029-988B-42

Query Match      21.9% Score 47; DB 6; Length 98;
Best Local Similarity 44.4%; Pred. No. 9.8;
Matches 8; Conservative 4; Mismatches 4; Indels 2; Gaps 1;

OY      6 KKDPSRKKEW--WENAGN 21
DB      38 RQAPGQRLEWGMWGNAGN 55

RESULT 9
US-10-032-037B-42

; Sequence 42, Application US/10032037B
; GENERAL INFORMATION:
; APPLICANT: Bio-Technology General Corp.
; TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
; TITLE OF INVENTION: MOIETIES, ANTIBODIES TO SUCH EPITOPES, AND USES THEREOF
; FILE REFERENCE: 10793/44
; CURRENT APPLICATION NUMBER: US/10/032,037B
; CURRENT FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/258,948
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 204
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 42
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-032-037B-42

Query Match      21.9% Score 47; DB 6; Length 98;
Best Local Similarity 44.4%; Pred. No. 9.8;
Matches 8; Conservative 4; Mismatches 4; Indels 2; Gaps 1;

OY      6 KKDPSRKKEW--WENAGN 21
DB      38 RQAPGQRLEWGMWGNAGN 55

RESULT 10
US-09-290-586A-19
; Sequence 19, Application US/09290586A
; GENERAL INFORMATION:
; APPLICANT: SANTAMARIA, Ignacio
; APPLICANT: VELASCO, Gloria
; APPLICANT: CAZORLA, Maite
; APPLICANT: FUEYO, Antonio
; APPLICANT: CAMPO, Elias
; APPLICANT: LOPEZ-OTIN, Carlos
; APPLICANT: AKOI, Takaoori
; APPLICANT: IMATA, Kazushi
; TITLE OF INVENTION: NOVEL HUMAN CATHEPSIN L2 PROTEIN, GENE ENCODING SAID
; TITLE OF INVENTION: PROTEIN AND USE THEREOF
; FILE REFERENCE: 99-410A/MMC/01332
; CURRENT APPLICATION NUMBER: US/09/290,586A
; CURRENT FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: JP 10-172147
; PRIOR FILING DATE: 1998-06-05
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Human
US-09-290-586A-19

Query Match      21.4% Score 46; DB 5; Length 329;
Best Local Similarity 52.9%; Pred. No. 45;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

OY      12 KKEWMENAGNKIYTMAL 28
DB      295 KNSWGNWGNKGYILMA 311

RESULT 11
US-10-287-971-18
; Sequence 18, Application US/10287971
; GENERAL INFORMATION:
; APPLICANT: Alsobrook, et al
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
; FILE REFERENCE: 21402-480A
; CURRENT APPLICATION NUMBER: US/10/287,971
; CURRENT FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: 09/997,425
```

```

; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 10/035,568
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: 60/338,626
; PRIOR FILING DATE: 2001-11-05
; PRIOR APPLICATION NUMBER: 60/401,479
; PRIOR FILING DATE: 2002-08-06
; PRIOR APPLICATION NUMBER: 60/333,072
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: 60/348,283
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: 60/393,262
; PRIOR FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: 60/406,181
; PRIOR FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 397
; SOFTWARE: CuroSeqList version 0.1
; SEQ ID NO: 18
; LENGTH: 1101
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-287-971-18
```

Query Match 21.4%; Score 46; DB 6; Length 1101;
Best Local Similarity 45.5%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Indels 0; Gaps 0;

Qy 8 DSRKKEWEN 18
Db 97 EPRDKKMW 107

```

RESULT 12
US-10-273-573-9476
; Sequence 9476, Application US/10273573
; GENERAL INFORMATION:
; APPLICANT: HySeq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: US/10/273,573
; CURRENT FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO: 9476
; LENGTH: 1839
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1331)..(1383)
; OTHER INFORMATION: Receptor tyrosine kinase class V proteins domain identified
; OTHER INFORMATION: by eMATRIX, accession number BL00790B, p-value=1.000e-40, raw scc
; OTHER INFORMATION: of 21.59
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1299)..(1458)
; OTHER INFORMATION: Ephrin receptor ligand binding domain identified by Pfam.
; OTHER INFORMATION: accession name EPH_1bd, E-value=1.4e-84, Pfam score of 294.4
; NAME/KEY: misc_feature
; LOCATION: (1)..(1839)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-273-573-9476
```

Query Match 21.2%; Score 45.5; DB 6; Length 1839;
Best Local Similarity 32.4%; Pred. No. 3.1e+02;
Matches 12; Conservative 9; Mismatches 9; Indels 7; Gaps 2;

Qy 10 SRKKEWENAGN-KIYTMADKTIS-----KLMTEY 39

Db 1406 TKSAFSEAPLYKVTIAADEFSQVDFGRLLAKDY 1442

```

RESULT 13
US-10-331-496A-60
; Sequence 60, Application US/10331496A
; GENERAL INFORMATION:
; APPLICANT: FRANTZ, GRETCHEN
; APPLICANT: HILLAN, KENNETH J.
; APPLICANT: PHILLIPS, HEIDI S.
; APPLICANT: POLAKIS, PAUL
; APPLICANT: SMITH, VICTORIA
; APPLICANT: SPENCER, SUSAN D.
; APPLICANT: WILLIAMS, P. MICKEY
; APPLICANT: WU, THOMAS D.
; APPLICANT: ZHANG, ZEMIN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF TUMOR
; FILE REFERENCE: P5014R1-PCT
; CURRENT APPLICATION NUMBER: US/10/331,496A
; CURRENT FILING DATE: 2002-12-30
; PRIOR APPLICATION NUMBER: US 60/345,444
; PRIOR FILING DATE: 2002-01-02
; PRIOR APPLICATION NUMBER: US 60/351,885
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: US 60/360,066
; PRIOR FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: US 60/362,004
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/366,869
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: US 60/366,284
; PRIOR FILING DATE: 2002-03-21
; PRIOR APPLICATION NUMBER: US 60/368,679
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 60/404,809
; PRIOR FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: US 60/405,645
; PRIOR FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 95
; SEQ ID NO: 60
; LENGTH: 606
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-331-496A-60
```

Query Match 20.7%; Score 44.5; DB 6; Length 606;
Best Local Similarity 41.0%; Pred. No. 1.3e+02;
Matches 16; Conservative 3; Mismatches 13; Indels 7; Gaps 3;

Qy 3 KVEKKDPSRKKEWENAGNKIYTMAD-KTISKLMTEYK 40
Db 462 KVQKKDP---KEM---AAQYREAMEADIKAAAEAAERK 494

```

RESULT 14
US-10-029-988B-41
; Sequence 41, Application US/10029988B
; GENERAL INFORMATION:
; APPLICANT: Bio-Technology General Corp.
; TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
; TITLE OF INVENTION: MOIETIES, ANTIBODIES TO SUCH EPITOPES, AND USES THEREOF
; FILE REFERENCE: 10793/46
; CURRENT APPLICATION NUMBER: US/10/029,988B
; CURRENT FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/258,948
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 204
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 41
; LENGTH: 98
; TYPE: PRT
```

; ORGANISM: Homo sapiens
US-10-029-988B-41

Query Match 20.5%; Score 44; DB 6; Length 98;
Best Local Similarity 44.4%; Pred. No. 23;
Matches 8; Conservative 4; Mismatches 4; Indels 2; Gaps 1;

OY 6 KKDPGRKKEM--WENAGN 21
::|::|::|::|
Db 38 ROAFQGRLEWMMGMWAGN 55

RESULT 15

US-10-032-037B-41
; Sequence 41, Application US/10032037B
; GENERAL INFORMATION:
; APPLICANT: Bio-Technology General Corp.
; TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
; FILE REFERENCE: 10793/44
; CURRENT APPLICATION NUMBER: US/10/032,037B
; CURRENT FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/258,948
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 204
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 41
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-032-037B-41

Query Match 20.5%; Score 44; DB 6; Length 98;
Best Local Similarity 44.4%; Pred. No. 23;
Matches 8; Conservative 4; Mismatches 4; Indels 2; Gaps 1;

OY 6 KKDPGRKKEM--WENAGN 21
::|::|::|::|
Db 38 ROAFQGRLEWMMGMWAGN 55

Search completed: July 24, 2003, 12:16:54
Job time : 2.3713 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 11:48:38 / Search time 3.27965 Seconds
(without alignments)
1172.914 Million cell updates/sec

Title: US-09-991-681-30

Perfect score: 215

Sequence: 1 SPKVEKKDPSRRKEMWENAGNKIYTMADKTISKLMTEYK 40

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	64	29.8	621	1 S59632	endo-1,4-beta-xylo
2	63.5	29.5	604	2 T41249	DEAD box ATP-depen
3	59.5	27.7	388	2 D71804	probable rod-shape
4	59	27.4	277	2 AB3084	oxidoreductase yaf
5	59	27.4	277	2 G98202	probable oxidoredu
6	58.5	27.2	104	2 AE9407	signal recognition
7	57.5	26.7	921	2 AE1698	isoencyl-tRNA syn
8	56.5	26.3	158	2 F90958	hypothetical prote
9	56.5	26.3	201	2 H85806	hypothetical prote
10	56.5	26.3	388	2 H64714	hypothetical prote
11	56	26.0	304	1 C84430	cell division prot
12	56	26.0	450	1 S73419	probable acid phos
13	55	25.6	177	2 T03232	signal recognition
14	54	25.1	387	2 S64082	protein BKG6 - com
15	54	25.1	1048	2 T04172	probable membrane
16	53.5	24.9	490	2 D75452	phosphoenolpyruvat
17	53.5	24.9	884	2 S73302	preprotein translo
18	53	24.7	259	2 T23782	hypothetical prote
19	53	24.7	410	2 T06665	hypothetical prote
20	53	24.7	425	2 B64316	restriction modifi
21	53	24.7	446	2 C64205	signal recognition
22	53	24.7	462	2 A70518	probable lipd prot
23	53	24.7	729	2 T16522	hypothetical prote
24	52.5	24.4	1289	2 B72354	conserved hypothet
25	52	24.2	277	2 T05658	hypothetical prote
26	52	24.2	433	2 T05693	hypothetical prote
27	52	24.2	599	2 S55363	maltase-like prote
28	52	24.2	840	2 C70131	leucine-tRNA ligas
29	52	24.2	1876	2 E97944	zinc metalloprotei

30	52	24.2	1984	2 A44396	P-type cation tran
31	52	24.2	26926	1 I38344	titin, cardiac mus
32	51.5	24.0	318	2 T13015	phosphoprotein pho
33	51.5	24.0	338	2 A82890	hypothetical prote
34	51.5	24.0	921	2 AC1327	isoencyl-tRNA syn
35	51	23.7	149	2 B96965	hypothetical prote
36	51	23.7	205	2 AE2442	hypothetical prote
37	51	23.7	230	2 C64396	precortin-2 methyl
38	51	23.7	238	2 A86160	hypothetical prote
39	51	23.7	342	2 E70109	hypothetical prote
40	51	23.7	547	2 A56575	putf-specific nucl
41	51	23.7	572	2 S73730	MG307 homolog H08
42	51	23.7	727	2 T24284	hypothetical prote
43	51	23.7	802	2 I39665	penicillin amidase
44	51	23.7	1017	2 T48452	hypothetical prote
45	51	23.7	1345	2 H90975	hypothetical prote

ALIGNMENTS

RESULT 1

S59632 endo-1,4-beta-xylinase (EC 3.2.1.8) B precursor - Cellvibrio mixtus

C:Species: Cellvibrio mixtus

C:Date: 15-Feb-1996 #sequence_revision 11-Apr-1997 #text_change 18-Jun-1999

C:Accession: S59632 #S2742

R:Millward-Sadler, S.J.; Davidson, K.; Hazlewood, G.P.; Black, G.W.; Gilbert, H.J.; Clax

Biochem. J. 312, 39-48, 1995

A:Title: Novel cellulose-binding domains, NodB homologues and conserved modular architec

A:Reference number: S59631; MUID:96077124; PMID:7492333

A:Accession: S59632

A:Molecule type: DNA

A:Residues: 1-621 <ML>

A:Cross-references: EMBL:Z48926; NID:9757808; PIDN:CAA68762.1; PID:9757809

C:Genetics:

A:Gene: xynB

C:Function:

A:Description: catalyzes the hydrolysis of 1,4-beta-xylosidic linkages in xylans

A:Pathway: xylan degradation

C:Superfamily: Pseudomonas endo-1,4-beta-xylinase F; Streptomyces endo-1,4-beta-xylinase

C:Keywords: glycosidase; hydrolase; polysaccharide degradation

F:1-19/Domains: signal sequence #status predicted <SIG>

F:20-621/Product: endo-1,4-beta-xylinase B #status predicted <MAT>

F:302-615/Domains: Streptomyces endo-1,4-beta-xylinase A homology <SXX>

F:403,516/Active site: Glu #status predicted

Query Match 29.8%; Score 64; DB 1; Length 621;

Best Local Similarity 43.2%; Pred. No. 3.7;

Matches 16; Conservative 5; Mismatches 16; Indels 0; Gaps 0;

Qy 2 PKVEKKDPSRRKEMWENAGNKIYTMADKTISKLMTE 38

Db 175 PIVQIKGSGYSGWGCWAGNELFTAGEDATISCTVTE 211

RESULT 2

T41249 DEAD box ATP-dependent RNA helicase - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Nov-2000

C:Accession: T41249

R:Seeger, K.; Harris, D.; Lyne, M.; Rajandream, M.A.; Barrett, B.G.

submitted to the EMBL Data Library, September 1998

A:Reference number: Z21981

A:Accession: T41249

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-604 <SER>

A:Cross-references: EMBL:AL031545; PIDN:CAA20842.1; GSPDB:GN00066; SPDB:SPCC285.03

A:Experimental source: strain 972h-; cosmid c285

D.: Jones, L.M.: Karet, U.
 Science 294, 849-852, 2001
 A:Authors: Krett, J.; Kuhn, M.; Kunst, F.; Kurapatk, G.; Madueno, E.; Maitouram, A.; Me
 ok, C.; Schuener, T.; Smoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,
 A:Title: Comparative genomics of *Listeria* species.
 A:Reference number: AB1077; MUID:21537279; PMID:11679669
 A:Accession: AE1698
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-921 <GLA>
 A:Cto8a- references: GS:AL592022; PIDN:CAC97357.1; PID:G16414641; GSPDB:GN00178
 A:Experimental source: strain Cl1p11262
 C:Genetics:
 A:Gene: ileS
 C:Superfamily: isoleucine-tRNA ligase

```

Query Match      26.7%; Score 57.5; DB 2; Length 921;
Best Local Similarity 26.2%; Pred. No. 36;
Matches 16; Conservative 7; Mismatches 9; Indels 29; Gaps 3

QY      9 PSRKKEW---WE-----NAGNKLYTM-----AADKTIISKLTMEY 39
      |::| | | | | | | | | | | | | | | | | | | | | |
Db      21 PNKEPEWQAKMBEEKLYEKIOEKONGRAYVILHDGPPYANGSELHNGHALNKTIKDIIVRY 80
      | 40 K 40
      | 81 K 81
Db

```

RESULT 8
P90958
hypothetical protein EC62638 [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509582)
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C:Accession: F90958
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G. G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
Nucleic Acids Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic
A:Reference number: A39629; MUID:21156231; PMID:11258796
A:Accession: F90958
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-158 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA836061.1; PID:G13362106; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RIMD 0509582
C:GeneticCB:
;Gene: EC62638

[illegible]

```

RESULT 9
H85806
hypothetical protein Z2984 [imported] - Escherichia coli (strain O157:H7, substrain EDL_57)
C:/Species: Escherichia coli
C:/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:/Accession: H85806
R:/Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,
Miller, L.; Grobecek, E.J.; Davis, N.W.; Llm, A.; Diallanta, E.; Potamousis, K.; Apodaca
Nature 409, 522-533, 2001
A:/Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:/Reference number: A85480; MUID:21074935; PMID:11206551
A:/Accession: H85806
A:/Status: preliminary
A:/Molecule type: DNA
A:/Residues: 1-201 <STO>

```

A;Cross-reference: GB:AE005174; NID:gl251596; PIDN:AA656916.1; GSPDB:GN00145; UMGF:Z239
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z2394

26.3%	Score 56.5;	DB 2;	length 201;
Best Local Similarity	35.7%;	Pred. No. 9.9;	
Matches	15;	Conservative	7; Mismatches 13; Indels 7; Gaps 2
Qy	5	EKKDPSRK-KEMWENA-----GNKITYMAADKITSIKMETY	39
		:	:: :::
Db	52	EKKQDRNRFLEFWMLKLANEMTYNGKILGKKAKKLANSKITNKF	93

RESULT 10
H64714
cell division protein - *Helicobacter pylori* (strain 26695)
C:Species: *Helicobacter pylori*
C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999
C:Accession: H64714
R:ToMb, J.F. ; White, O. ; Kerlavage, A.R. ; Clayton, R.A. ; Sutton, G.G. ; Fleischmann, R.D. ; Peterson, S. ; Loftus, B. ; Richardson, D. ; Dodson, R. ; Khakel, H.G. ; Glodek, A. ; McMeekin, J.D. ; Kelley, J.M. ; Cotton, M.D. ; Weidman, J.M. ; Fujii, C. ; Bowman, C. ; Watney, L. ; NATURE 368, 539-547, 1997
A:Authors: Wallin, E. ; Hayes, W.S. ; Borodovsky, M. ; Karpk, P.D. ; Smith, H.O. ; Fraser, C.
A:Title: The complete genome sequence of the gastric pathogen *Helicobacter pylori*.
A:Reference number: A64520; MUID:97394467; PMID:9252185
A:Accession: H64714
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-388 <TOM>
A:Cross-references: GB:AE000654; GB:AE000511; NID:g2314743; PIDN:AAD08600.1; PID:g2314743
!superfamily: rod shape-determining protein

Query Match	26.3%	Score	56.5	DB	2	Length	388
Best Local Similarity	34.3%	Pred.	No.	20			
Matches	12	Conservative	5	Mismatches	13	Indels	5
				Gaps	1		
QY	11	RKKEWMENAGKITYMAADK	-----	TSKMTTEK	40		
			:	:	:		
DB	217	RLKLMSTLONSLEFLDPLRANLRISDDEST	251				

RESULT 11
 C64430
 Probable acid phosphatase (EC 3.1.3.2) At2g01880 precursor [similarity] - Arabidopsis th
 N/Alternate names: purple acid phosphatase
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Nov-2001
 C/Accession: C64430
 R/In, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 W.; Koo, H.; Moffatt, K.S.; Cronin, L.A.; Shen, M.; Vanhaken, S.E.; Umayam, L.;
 Eusner, D.; Niemman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
 N. Nature 402, 761-768, 1999
 A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A/Reference number: A84420; MUID:20083487; PMID:10617197
 A/Accession: C64430
 A/Molecule type: DNA
 A/Residues: 1-304 <STO>
 A/Cross-references: GB:EM002093; NID:g4522007; PIDD:AMD21780.1; GSPDB:GN00139
 C/Genetics:
 A/Gene: At2g01880
 A/Map position: 2
 C/Superfamily: tartrate-resistant acid phosphatase; phosphoserine/threonine phosphatase
 C/Keywords: iron; metalloprotein; phosphoric monoester hydrolase
 E:1-21/Domain: signal sequence #status predicted <Sig>
 E:22-304/Product: acid phosphatase #status predicted <Mar>
 F:42-121/Domain: phosphoserine/threonine phosphatase core homology <PCC>
 F:48,81,84,226/Binding site: iron (Asp, Asp, Tyr, His) #status predicted
 F:81,119,189,224/Binding site: iron (Asp, Asn, His, His) #status predicted
 F:120,198/Active site: His #status predicted

Query Match 26.0%; Score 56; DB 1; Length 304;

Accession: S64082

Job time : 6.27965 secs

Accession: S64082

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 11:46:57 ; Search time 1.81552 Seconds

(without alignments)
1036.105 Million cell updates/sec

Title: US-09-991-681-30

Perfect score: 215
Sequence: 1 SPKYEKDPSSRKEMWENAGNKIYTMADKTIKLTMEYK 40

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	59.5	27.7	388	1	FTSW_HELPU
2	58.5	27.2	104	1	SR19_ARCRU
3	56.5	26.3	388	1	FTSW_HELPU
4	56	26.0	450	1	SR54_MYCPN
5	55.5	25.8	103	1	SR19_METKA
6	54	25.1	321	1	CS37_CANAL
7	54	25.1	387	1	YGH5_YEAST
8	53.5	24.9	466	1	PECK_DEIRA
9	53.5	24.9	884	1	SBGA_PORPU
10	53	24.7	343	1	Y130_METJA
11	53	24.7	446	1	SR54_MYCGB
12	52	24.2	468	1	SG3_HUMAN
13	52	24.2	840	1	SVL_BORBU
14	52	24.2	1956	1	ATXI_PLATA
15	51.5	24.0	322	1	PP16_ARATH
16	51.5	24.0	747	1	ORPB_HUMAN
17	51	23.7	230	1	Y771_METJA
18	51	23.7	230	1	MAT4_BACSH
19	51	23.7	572	1	EX42_DROME
20	51	23.7	605	1	Y337_MYCPN
21	51	23.7	802	1	DME_ARATH
22	51	23.7	1729	1	DME_ARATH
23	51	23.7	2660	1	YEEJ_ECO57
24	51	23.7	4543	1	LRP1_CHICK
25	51	23.7	4543	1	LRP1_CHICK
26	51	23.7	4543	1	LRP1_CHICK
27	50.5	23.5	445	1	AS3A_MOUSE
28	50.5	23.5	745	1	Y700_YEAST
29	50	23.3	208	1	GTHB_ARATH
30	50	23.3	209	1	GTHB_ARATH
31	50	23.3	359	1	SUR6_FUGRU
32	50	23.3	505	1	VLA_HPV07
33	50	23.3	600	1	SPO8_YEAST

34	50	23.3	859	1	ENV_EIAV3
35	50	23.3	1070	1	PVDR_PLAVS
36	50	23.3	1118	1	UBP8_HUMAN
37	50	23.3	1409	1	AEX3_CABEL
38	50	23.3	1469	1	DE27_CABEL
39	50	23.3	2358	1	YEEJ_ECOLI
40	49.5	23.0	178	1	EFAB_CHICK
41	49.5	23.0	235	1	YOF9_CABEL
42	49.5	23.0	310	1	PP12_TOBAC
43	49	22.8	368	1	RA11_HUMAN
44	49	22.8	411	1	YG41_YEAST
45	49	22.8	427	1	GGNT_BOVIN

ALIGNMENTS

RESULT 1	ID	FTSW_HELPU	STANDARD	PRT	388 AA.
AC	092348				
DT	16-OCT-2001	(Rel. 40, Created)			
DT	16-OCT-2001	(Rel. 40, Last sequence update)			
DT	16-OCT-2001	(Rel. 40, Last annotation update)			
DE	Probable cell division protein ftsw.				
GN	FTSW OR JHP1468.				
OS	Helicobacter pylori J99 (Campylobacter pylori J99).				
OC	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;				
OC	Helicobacteraceae; Helicobacter.				
OK	NCBI_TaxID=85963;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=99120557; Pubmed=9923682;				
RA	Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,				
RA	Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,				
RA	Tummino P.J., Caruso A., Ulla-Nickelsen M., Mills D.W., Ives C.,				
RA	Gibson R., Werberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,				
RA	Trust T.J.;				
RT	"Genomic sequence comparison of two unrelated isolates of the human				
RT	gastric pathogen Helicobacter pylori.";				
RL	Nature 397:176-180(1999).				
CC	- FUNCTION: THIS IS A SEPTUM-PEPTIDOGLYCAN BIOSYNTHETIC PROTEIN,				
CC	INVOLVED IN CELL WALL FORMATION. PLAYS A ROLE IN THE STABILIZATION				
CC	OF THE FTSG RING DURING CELL DIVISION (BY SIMILARITY).				
CC	- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane				
CC	(by similarity).				
CC	- SIMILARITY: BELONGS TO THE FTSW/RODA/SPOVE FAMILY.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL: AB001568; AAD07042.1; -				
DR	PIR: D71804; D71804.				
DR	InterPro: IPR001182; Cell cycle.				
DR	PIfam: PF01098; FTSW_RODA_SPOVE; 1.				
DR	PROSITE: PS00428; FTSW_RODA_SPOVE; 1.				
KW	Peptidoglycan synthesis; Cell wall; Cell division; Cell shape;				
KW	Transmembrane; Inner membrane; Complete proteome.				
FT	TRANSMEM	7	27		POTENTIAL.
FT	TRANSMEM	39	59		POTENTIAL.
FT	TRANSMEM	71	91		POTENTIAL.
FT	TRANSMEM	106	126		POTENTIAL.
FT	TRANSMEM	145	165		POTENTIAL.
FT	TRANSMEM	177	197		POTENTIAL.
FT	TRANSMEM	262	282		POTENTIAL.
FT	TRANSMEM	295	315		POTENTIAL.
FT	TRANSMEM	328	348		POTENTIAL.
FT	TRANSMEM	359	379		POTENTIAL.

SQ SEQUENCE 388 AA; 42540 MW; 51E4277A910B286D CRC64;
 Query Match 27.7%; Score 59.5; DB 1; Length 388;
 Best Local Similarity 37.1%; Pred. No. 3.2;
 Matches 13; Conservative 4; Mismatches 13; Indels 5; Gaps 1;

QY 11 RKKEWENAGNKITYTMAADK-----TISTLMTEYK 40
 DB 217 RLKLWMSNLQNSLFTLLPKLANALRISDLPSSTQ 251

RESULT 2
 SRP19 ARCFU STANDARD; PRT; 104 AA.
 AC 029010;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Signal recognition particle 19 kDa protein (SRP19).
 GN SRP19 OR AF1258.
 OS Archaeoglobus fulgidus.
 OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
 OC Archaeoglobaceae; Archaeoglobus.
 OX NCBI_TaxId=2234;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE=98049343; PubMed=9389475;
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Peterson K.E.,
 RA Ketchum K.A., Dodson R.J., Gwin M., Hickey E.K., Peterson J.D.,
 RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyriades N.C.,
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 RA Kirschner E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
 RA Cotton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.;
 RA "The complete genome sequence of the hyperthermophilic, sulphate-
 RT reducing archaeon Archaeoglobus fulgidus.";
 RT Nature 390:364-370(1997).
 RN [2]
 RP CHARACTERIZATION.
 RX MEDLINE=20150251; PubMed=10684931;
 RA Bhuiyan S.H., Gowda K., Hotkezaka H., Zwiab C.;
 RT "Assembly of archaeal signal recognition particle from recombinant
 RT components.";
 RT Nucleic Acids Res. 28:1365-1373(2000).
 RN [3]
 RP CHARACTERIZATION.
 RX MEDLINE=20496765; PubMed=11041851;
 RA Diener J.L., Wilson C.;
 RT "Role of SRP19 in assembly of the Archaeoglobus fulgidus signal
 RT recognition particle.";
 RT Biochemistry 39:12862-12874(2000).
 RN [4]
 RP STRUCTURE BY NMR.
 RX MEDLINE=21914566; PubMed=11916385;
 RA Pakhomova O.N., Deep S., Huang O., Zwiab C., Hinck A.P.;
 RT "Solution structure of protein SRP19 of Archaeoglobus fulgidus signal
 RT recognition particle.";
 RT J. Mol. Biol. 317:145-158(2002).
 CC -1- FUNCTION: SIGNAL-RECOGNITION-PARTICLE ASSEMBLY, BINDS DIRECTLY
 CC TO 7S RNA AND MEDIATES BINDING OF THE 54 KDa SUBUNIT OF THE SRP.
 CC -1- SUBUNIT: ARCHAEL SIGNAL RECOGNITION PARTICLE CONSISTS OF A 7S RNA
 CC MOLECULE OF 300 NUCLEOTIDES AND TWO PROTEIN SUBUNITS: SRP54 AND
 CC SRP19.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE SRP19 FAMILY.
 CC -1- SIMILARITY: BELONGS TO THE SRP19 FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL, AE001017; AAB89988.1; -
 DR PIR, A69407; A69407.
 DR PDB, 1KVN; 20-MAR-02.
 DR PDB, 1KVV; 20-MAR-02.
 DR TIGR, AF1258; -
 DR HAMAP, MF_00305; -; 1.
 DR InterPro, IPR002778; SRP19.
 DR Pfam, PF01922; SRP19; 1.
 DR ProDom, PD006609; SRP19; 1.
 DR Signal recognition particle; RNA-binding; Ribonucleoprotein;
 KW Complete proteome; 3D-structure.
 SQ SEQUENCE 104 AA; 12405 MW; 72D5DDAB84E89E64 CRC64;

Query Match 27.2%; Score 59.5; DB 1; Length 104;
 Best Local Similarity 44.1%; Pred. No. 1.1;
 Matches 15; Conservative 1; Mismatches 13; Indels 5; Gaps 2;

QY 5 EKQPSRKKEWENAGNKITYTMAADKTTISKLMTE 38
 DB 52 EKTYP--KSWMBEGRVVEKRGTKT--KLMI 80

RESULT 3
 FTSM HELPY STANDARD; PRT; 388 AA.
 AC P56096;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable cell division protein ftsw.
 GN FTSM OR HP1560.
 OS Helicobacter pylori (Campylobacter pylori).
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
 OC Helicobacteraceae; Helicobacter.
 OX NCBI_TaxId=210;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=26695 / ATCC 700392;
 RX MEDLINE=97934467; PubMed=9252185;
 RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
 RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
 RA Nelson K., Quackenbush J., Zhou L., Kirschner E.F., Peterson S.,
 RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
 RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
 RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
 RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Wathey L., Wallin E.,
 RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
 RA Venter J.C.;
 RT "The complete genome sequence of the gastric pathogen Helicobacter
 RT pylori.";
 RT Nature 388:539-547(1997).
 CC -1- FUNCTION: THIS IS A SEPTUM-PEPTIDOGLYCAN BIOSYNTHETIC PROTEIN,
 CC INVOLVED IN CELL WALL FORMATION. PLAYS A ROLE IN THE STABILIZATION
 CC OF THE FTSG RING DURING CELL DIVISION (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
 CC (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE FTSM/RODA/SPOVE FAMILY.
 CC -1- SIMILARITY: BELONGS TO THE FTSM/RODA/SPOVE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL, AE000654; AAD08600.1; -

```

DR PIR; H64714; H64714.
DR TIGR; HPL1560; -.
DR InterPro; IPR001182; Cell cycle.
DR Pfam; PF01098; FTSW_R0D4_SPOVE; 1.
DR PROSITE; PS00428; FTSW_R0D4_SPOVE; 1.
DR Peptidoglycan synthesis; Cell wall; Cell division; Cell shape;
KW Transmembrane; Inner membrane; Complete proteome.
FT TRANSMEM 7 27 POTENTIAL.
FT TRANSMEM 39 59 POTENTIAL.
FT TRANSMEM 71 91 POTENTIAL.
FT TRANSMEM 106 126 POTENTIAL.
FT TRANSMEM 145 165 POTENTIAL.
FT TRANSMEM 177 197 POTENTIAL.
FT TRANSMEM 262 282 POTENTIAL.
FT TRANSMEM 295 315 POTENTIAL.
FT TRANSMEM 328 348 POTENTIAL.
FT TRANSMEM 359 379 POTENTIAL.
SQ SEQUENCE 388 AA; 42478 MW; 77195EA8052BPF1 CRC64;

Query Match 26.3%; Score 56.5; DB 1; Length 388;
Best Local Similarity 34.3%; Pred. No. 7.7;
Matches 12; Conservative 5; Mismatches 13; Indels 5; Gaps 1;

Oy 11 RKKEWENAGNKITYTMADK-----TISKLMTEYK 40
Db 217 RLKLMWSNMQNSLFTLPDLRLANALRIDLPESYQ 251

RESULT 4
SR54_MYCPN STANDARD; PRT; 450 AA.
ID SR54_MYCPN
AC P75054;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Signal recognition particle protein (fifty-four homolog).
GN Ffh OR MP061 OR MP093.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=9710585; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirk E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -1- FUNCTION: NECESSARY FOR EFFICIENT EXPORT OF EXTRA-CYTOPLASMIC
CC PROTEINS. BINDS TO THE SIGNAL SEQUENCE WHEN IT EMERGES FROM THE
CC RIBOSOMES (BY SIMILARITY).
CC -1- SUBUNIT: SIGNAL RECOGNITION PARTICLE CONSISTS OF A SMALL
CC CYTOPLASMIC RNA (SC-RNA) MOLECULE AND PROTEIN Ffh (BY SIMILARITY).
CC -1- DOMAIN: THE PROTEIN HAS A TWO DOMAIN STRUCTURE: THE G-DOMAIN
CC BINDS GTP. THE M-DOMAIN BINDS THE RNA AND ALSO BINDS THE SIGNAL
CC SEQUENCE.
CC -1- SIMILARITY: BELONGS TO THE SRP FAMILY OF GTP-BINDING PROTEINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; AE000011; AAB35741.1; -.
DR PIR; S73419; S73419.
DR HSSP; O07347; 1FPH.
DR InterPro; IPR003593; AAA_ArPase.
DR InterPro; IPR008977; SRP54.
DR InterPro; IPR004125; SRP54_SBP.

```

```

DR InterPro; IPR004780; SRP_sub.
DR Pfam; PF00448; SRP54_1.
DR Pfam; PF02881; SRP54_N; 1.
DR Pfam; PF02978; SRP_SBP; 1.
DR ProDom; PD000819; SRP54; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRfams; TIGR00959; ffh; 1.
DR PROSITE; PS00300; SRP54; 1.
KW Signal recognition particle; GTP-binding; RNA-binding;
KW Complete proteome.
FT DOMAIN 1 293 G-DOMAIN.
FT DOMAIN 294 450 M-DOMAIN.
FT NP_BIND 106 113 GTP (BY SIMILARITY).
FT NP_BIND 188 192 GTP (BY SIMILARITY).
FT NP_BIND 246 249 GTP (BY SIMILARITY).
SQ SEQUENCE 450 AA; 50141 MW; 7EFA42ECAA4FE679 CRC64;

Query Match 26.0%; Score 56; DB 1; Length 450;
Best Local Similarity 34.2%; Pred. No. 10;
Matches 13; Conservative 9; Mismatches 10; Indels 6; Gaps 1;

Oy 2 PKVEKQSPRKKEWENAGNKITYTMADKTIISKLMTEY 39
Db 387 PKLINRDPSPRRKRIIKSGRRM-----DELNKLKKEW 418

RESULT 5
SR19_METKA STANDARD; PRT; 103 AA.
ID SR19_METKA
AC Q8TV49;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Signal recognition particle 19 kDa protein (SRP19).
GN SRP19 OR MK1551.
OS Methanopyrus kandleri.
OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
OC Methanopyrus.
OX NCBI_TaxID=2320;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=AV19 / DSM 6324 / JCM 9639;
RX MEDLINE=21927647; PubMed=11930014;
RA Slesarev A.I., Mezhevaya K.V., Makarova K.S., Polushin N.N.,
RA Shcherbina O.V., Shakhova V.V., Belova G.I., Arvid L.,
RA Natile D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
RA Matykh A.G., Koonin E.V., Kozlovskiy S.A.;
RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19
RT and monophyly of archaeal methanogens.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
CC -1- FUNCTION: Signal-recognition-particle assembly, binds directly to
CC 7S RNA and mediates binding of the 54 kDa subunit of the SRP (by
CC similarity).
CC -1- SUBUNIT: Archaeal signal recognition particle consists of a 7S RNA
CC molecule of 300 nucleotides and two protein subunits: SRP54 and
CC SRP19 (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SRP19 FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; AB010445; AAM02764.1; -.
DR HAMAP; MF_00305; -; 1
DR InterPro; IPR002778; SRP19.
DR Pfam; PF01922; SRP19; 1.
DR ProDom; PD006609; SRP19; 1.
KW Signal recognition particle; RNA-binding; Ribonucleoprotein;

```

[illegible]

```

NS YG075C Saccharomyces cerevisiae (Baker's Yeast) .
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=5288C;
EX MEDLINE=97435481; PubMed=9290212;
RA Rieger M., Brueckner M., Schaefer M., Mueller-Auer S.;
RT "Sequence analysis of 203 kilobases from Saccharomyces cerevisiae
chromosome VII";
RL Yeast 13:1077-1090 (1997) .
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC -----
CC DR EMBL; 272597; CA96780.1; -
CC DR PIR; S64082; S64082.
CC DR SGD; S0003043; MIPS2.
CC DR GO; GO:0005635; C:nuclear membrane; IDA.
CC DR GO; GO:0005816; C:spindle pole body; IDA.
CC DR GO; GO:0005200; F:structural constituent of cytoskeleton; IMP.
CC KW Hypothetical protein; Transmembrane.
CC FT TRANSMEM 311 327 POTENTIAL.
CC ST SEQUENCE 387 AA; 44585 MW; C88AFA665DEA834C CRC64;
CC -----
Oy Query Match 25.1%; Score 54; DB 1; Length 387;
Db Best Local Similarity 46.4%; Pred. No. 16;
Matches 13; Conservative 4; Mismatches 7; Indels 4; Gaps 2;
Oy 15 WVENAG--NKITYMAADTKISKLTMTXK 40
Db 347 WVENSGILSKIVWFFEDQ--DLETRYR 372
Oy |||||:|:|||||:|||||:
Db |||||:|:|||||:|||||:
RESULT 8
PCKC DEIRA STANDARD; PRT; 466 AA.
ID_PCKC DEIRA STANDARD; PRT; 466 AA.
AC QSRVPE;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phosphoenolpyruvate carboxykinase [KTP] (EC 4.1.1.49) (PEP
carboxykinase) (Phosphoenolpyruvate carboxylase) (PEPCK).
GN PCKA OR DR0977.
OS Deinococcus radiodurans.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
CC NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RP STRAIN=R1 / ATCC 13939 / DSM 20539 / NCIB 9279;
RC MEDLINE=20036896; PubMed=10567266;
RX White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
Dodson R.J., Helt D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
Vanatkevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,
Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
Kechum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
radiodurans R1";
RL Science 286:1571-1577 (1999) .
CC -1- CATALYTIC ACTIVITY: ATP + oxaloacetate = ADP + phosphoenolpyruvate
+ CO(2).
CC -1- PATHWAY: Rate-limiting gluconeogenic enzyme.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity) .

```

```

CC -1- SIMILARITY: Belongs to the phosphoenolpyruvate carboxykinase [ATP]
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE001950; AAF10554.1; ALT_INIT.
CC HSSP; P22259; IATL.
CC TIGR; DR0977; -.
CC HAMAP; MF_00453; acylal; 1.
CC InterPro; IPR001272; PEPCK_ATP.
CC Pfam; PF01293; PEPCK_ATP_1.
CC Prodom; PD004723; PEPCK_ATP_1.
CC TIGRFAMs; TIGR00224; pckA; 1.
CC PROSITE; PS00532; PEPCK_ATP_1.
CC Glucosyltransferase; Lyase; Decarboxylase; ATP-binding; Complete proteome.
CC NP BIND 237 244 ATP (BY SIMILARITY).
CC SEQUENCE 466 AA; 50951 MW; 22PDC1A23215942 CRC64;

Query Match 24.9%; Score 53.5; DB 1; Length 466;
Best Local Similarity 28.2%; Pred. No. 22;
Matches 11; Conservative 8; Mismatches 17; Indels 3; Gaps 1;

Qy 1 SPK---VEKDPKSKKEMWENAGNKIYTMADTKTSKLM 36
Db 62 SPKRFIVEDQGTRETVWEGFNQPIASAEVDFRLDKNV 100

RESULT 9
SECA_PORPU STANDARD; PRT; 884 AA.
ID SECA_PORPU
AC P51381;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Preprotein translocase seca subunit.
GN SECA.
OS Porphyra purpurea.
OC Chloroplast.
OC Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.
OX NCBI_TaxId=2787;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Avonport;
RA Reith M.E., Munholland J.;
RT "Complete nucleotide sequence of the Porphyra purpurea chloroplast
RT genome.";
RL Plant Mol. Biol. Rep. 13:333-335(1995).
CC -1- FUNCTION: INVOLVED IN PROTEIN EXPORT AND IMPORT. SECA HAS A
CC CENTRAL ROLE IN COUPLING THE HYDROLYSIS OF ATP TO THE TRANSFER OF
CC PRE-SECRETORY PROTEINS ACROSS THE MEMBRANE.
CC -1- SUBCELLULAR LOCATION: Chloroplast.
CC -1- SIMILARITY: BELONGS TO THE SECA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U38804; AAC08267.1; -.
CC PIR; S73302; S73302.
CC InterPro; IPR000185; SecA.
CC Pfam; PF01043; SecA_protein; 1.
CC PRINTS; PR00906; SECA.
CC TIGRFAMs; TIGR00963; seca; 1.

```

```

DR PROSITE; PS01312; SECA; 1.
KW Protein transport; ATP-binding; Chloroplast; Translocation; Transport.
KW NP BIND 98 105 ATP (BY SIMILARITY).
SQ SEQUENCE 884 AA; 101325 MW; A41584601259082B CRC64;

Query Match 24.9%; Score 53.5; DB 1; Length 884;
Best Local Similarity 28.1%; Pred. No. 43;
Matches 16; Conservative 9; Mismatches 7; Indels 25; Gaps 3;

Qy 9 PSRKEM-----WENAGNKIY-----TMAADKT--ISKLMTEYK 40
Db 394 PLRRKESDLVYSNRYKWEALADECYDMRAGRPTLVGTYSVKSSELSLTLEYK 450

RESULT 10
Y130_METUA STANDARD; PRT; 343 AA.
ID Y130_METUA
AC Q57594;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein M0130.
GN M0130.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxId=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kierlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reith C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Uteerback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenn H.-P., Fraser C.M., Smith H.O., Weese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.";
RL Science 273:1058-1073(1996).
RN [2]
RP REVISIONS.
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kierlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reith C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Uteerback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenn H.-P., Fraser C.M., Smith H.O., Weese C.R., Venter J.C.;
RL Submitted (FEB-1998) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: STRONG, TO M_JANNASCHII M01218 AND M01531.
CC -1- SIMILARITY: TO TYPE I RESTRICTION ENZYME SPECIFICITY PROTEIN (S
CC SUBUNIT).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U67470; AAB98112.1; -.
CC TIGR; M0130; -.
CC InterPro; IPR000055; Resc_mod_DNA.
CC Pfam; PF01420; Methylase_S_2_
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 343 AA; 39061 MW; 5184D6866274603 CRC64;

Query Match 24.7%; Score 53; DB 1; Length 343;

```

Best Local Similarity 29.5%; Pred. No. 19;
Matches 13; Conservative 8; Mismatches 9; Indels 14; Gaps 2;

QY 9 PRRKK-EMWENA-----GNKIYWAADTKSLKMT 38
Db 245 PSTRKSEYENGEMINWITPLDLSRLNEXIYIGSSERKTRIALE 288

RESULT 11

SR54_MYCGE STANDARD; PRT; 446 AA.
ID SR54_MYCGE
AC P47294;
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Signal recognition particle protein (fifty-four homologs).
GN F5H OR MG048.
OS Mycoplasma genitalium.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxId=2097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=96026346; PubMed=7569993;
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Furumasa J.L.,
RA Nguyen D.T., Uettermann T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-F., Dougherty B.A., Boer K.F., Hu P.-C., Lincer T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT "The minimal gene complement of Mycoplasma genitalium.";
RL Science 270:397-403 (1995).
CC -!- FUNCTION: NECESSARY FOR EFFICIENT EXPORT OF EXTRA-CYTOPLASMIC
CC PROTEINS. BINDS TO THE SIGNAL SEQUENCE WHEN IT EMERGES FROM THE
CC RIBOSOMES (BY SIMILARITY).
CC -!- SUBUNIT: SIGNAL RECOGNITION PARTICLE CONSISTS OF A SMALL
CC CYTOPLASMIC RNA (SC-RNA) MOLECULE AND PROTEIN F5H (BY SIMILARITY).
CC -!- DOMAIN: THE PROTEIN HAS A TWO DOMAIN STRUCTURE: THE G-DOMAIN
CC BINDS GTP; THE M-DOMAIN BINDS THE RNA AND ALSO BINDS THE SIGNAL
CC SEQUENCE.
CC -!- SIMILARITY: BELONGS TO THE SRP FAMILY OF GTP-BINDING PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U39684; AAC1264.1; -.
DR PIR: C64205; C64205.
DR HSSP: 007347; 1FFH.
DR TIGR: MG048; -.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR000897; SRP54.
DR InterPro: IPR004125; SRP54_SPB.
DR InterPro: IPR004780; SRP_SDB.
DR Pfam: PF00448; SRP54; 1.
DR Pfam: PF02881; SRP54_N; 1.
DR Pfam: PF02978; SRP_SPB; 1.
DR ProDom: PD000819; SRP54; 1.
DR SMART: SM00382; AAA; 1.
DR TIGRfam: TIGR00959; fth.1.
DR PROSITE: PS00300; SRP54; 1.
KW Signal recognition particle; GTP-binding; RNA-binding;
KW Complete proteome.
FT DOMAIN 1 293 G-DOMAIN.
FT DOMAIN 294 446 M-DOMAIN.
FT NP_BIND 106 113 GTP (BY SIMILARITY).
FT NP_BIND 188 192 GTP (BY SIMILARITY).
FT NP_BIND 246 249 GTP (BY SIMILARITY).
SQ SEQUENCE 446 AA; 50199 MW; 080F2255CB85E5AE CRC64;

Query Match 24.7%; Score 53; DB 1; Length 446;
Best Local Similarity 31.6%; Pred. No. 25;
Matches 12; Conservative 10; Mismatches 10; Indels 6; Gaps 1;

QY 2 PVEKKDPSRKKEMWENAGNKIYWAADTKSLKMT 39
Db 367 PRLINRDNRRKRRIRIKSGRCM-----DELNKLKMK 418

RESULT 12

SG3_HUMAN STANDARD; PRT; 468 AA.
ID SG3_HUMAN
AC Q8WMD2; Q96C83; Q96GE8; Q916G7;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Secretogranin III precursor (SgIII).
GN SCG3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary;
RA Song H., Peng Y., Huang Q., Dai M., Mao Y., Zhang Q., Mao M., Fu G.,
RA Luo M., Chen J., Hu R.;
RT "Human secretogranin III mRNA, complete cds.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RX MEDLINE=22093564; PubMed=12098761;
RA Rong Y.P., Liu F., Zeng L.C., Ma W.J., Wei D.Z., Han Z.G.;
RT "Cloning and characterization of a novel human secretory protein:
RL secretogranin III.";
RL Acta Biochim. Biophys. Sin. 34:411-417 (2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain, and Lung;
RX MEDLINE=22388257; PubMed=12477932;
RA Strussberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Nak S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marutina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smalhus D.E.,
RA Schnerich A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- SUBCELLULAR LOCATION: Secreted. Neuroendocrine and endocrine
CC secretory granules.
CC -!- TISSUE SPECIFICITY: Expressed in brain, heart, kidney, liver and
CC skeletal muscle.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

```

CC -----
DR EMBL; AF078851; AAD4483.1; -
DR EMBL; AF453583; AAL6743.1; -
DR EMBL; BC014539; AAH14539.1; -
DR EMBL; BC009511; AAH09511.1; ALT_INIT.
DR Genew; HGNC:11707; SGC3.
KM Signal; Cleavage on pair of basic residues; Polymorphism.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 468 SECRETORANIN III.
FT VARIANT 125 125 S -> N (in dbSNP:2305710).
FT CONFLICT 79 79 K -> R (IN REF. 1).
FT CONFLICT 167 167 A -> V (IN REF. 1 AND 2).
FT CONFLICT 272 274 EEL -> RDF (IN REF. 1).
SQ SEQUENCE 468 AA; 52977 MW; 633A87692A783808 CRC64;

Query Match
Best Local Similarity 32.6%; Score 52; DB 1; Length 468;
Matches 15; Conservative 8; Mismatches 9; Indels 14; Gaps 3;

QY 6 KKDPSRKKEWMEN-AG--NKIYTMAA-----DKTISKMT 37
DB 209 EEDPKPTSWTENQAGKIPKVTPLAALQDGLAKGENDETVSNTLT 254

RESULT 13
SYL_BORBU STANDARD; PRT; 840 AA.
AC 051267;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE leucyl-tRNA synthetase (EC 6.1.1.4) (leucine--tRNA ligase) (leuRS).
GN LEUS OR BB0251.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OK NCBI_TaxID=139;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685;
RA Frazer C.M., Casjens S., Huang W.W., Sutton G.G., Clayton R.A.,
RA Lachlagra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Kierlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA Van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weiman J.,
RA Utermbeck T., Matthey L., McDonald L., Artlich P., Bowman C.,
RA Gaitland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochete, Borrelia
RT burgdorferi."
RL Nature 390:580-586(1997).
CC -1- CATALYTIC ACTIVITY: ATP + L-leucine + tRNA(Leu) = AMP +
CC -1- diphosphate + L-leucyl-tRNA(Leu).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE001135; AAB91495.1; -
CC PIR; C70131; C70131.
CC TIGR; BB0251; -
CC HAMAP; MF_00049; -; 1.
CC InterPro; IPR002302; Leu-tRNA-synth.
CC InterPro; IPR002300; tRNA-synth.1a.
CC InterPro; IPR001412; tRNA-synth.1.
CC Pfam; PF00133; tRNA-synth.1; 1.

```

```

DR PRINTS; PR00985; TRNASYNTHLEU.
DR TIGRFAMs; TIGR00396; leuS_bact; 1.
DR PROSITE; PS00178; AA TRNA_LIGASE I; 1.
KM Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KM Complete proteome.
FT SITE 44 55 "HIGH" REGION.
FT SITE 617 621 "KMSK" REGION.
FT BINDING 620 620 ATP (BY SIMILARITY).
SQ SEQUENCE 840 AA; 98158 MW; 521C76BAF009593E CRC64;

Query Match
Best Local Similarity 40.6%; Score 52; DB 1; Length 840;
Matches 13; Conservative 6; Mismatches 7; Indels 6; Gaps 2;

QY 3 KVEKDPSSRKKEWMENGNKITYTMAADKTSK 34
DB 8 KIEKK---WQEFWDN--NKITYVEEDPSIFK 33

RESULT 14
ATX1_PLAFA STANDARD; PRT; 1956 AA.
AC 004956;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Probable cation-transporting ATPase 1 (EC 3.6.3.-).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OK NCBI_TaxID=5833;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=T9/96;
RX MEDLINE=93132070; PubMed=8421054;
RA Krishna S., Cowan G., Meade J.C., Wells R.A., Stringer J.R.,
RA Robson K.J.;
RT "A family of cation ATPase-like molecules from Plasmodium
RT falciparum."
RL J. Cell Biol. 120:385-398(1993).
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + phosphate.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: Belongs to the cation transport ATPases family (P-type
CC ATPases). Subfamily V.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X65738; CAA4646.1; -
CC InterPro; IPR001757; ATPase_E1-E2.
CC Pfam; PF00122; E1-E2_ATPase; 1.
CC TIGRFAMs; TIGR01494; ATPase_P-type; 7.
DR HydroLase; PS00154; ATPase_E1_E2; 1.
KM HydroLase; Transmembrane; Phosphorylation; Magnesium; ATP-binding.
FT DOMAIN 1 35 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 36 58 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 59 61 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 62 80 POTENTIAL.
FT DOMAIN 81 407 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 408 427 POTENTIAL.
FT DOMAIN 428 440 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 441 462 POTENTIAL.
FT DOMAIN 463 1818 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 1819 1837 POTENTIAL.
FT DOMAIN 1838 1845 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 1846 1863 POTENTIAL.
FT DOMAIN 1864 1881 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 1882 1905 POTENTIAL.
FT DOMAIN 1906 1928 EXTRACELLULAR (POTENTIAL).

```

RA Petteret A., Rajendream M.A., Lyn M., Benes V., Rechmann S.,
RA Borokva D., Raedcker H., Scharfe M., Grimm M., Loehnert T.-H.,
RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fatmann B., Grandtath K., Dauner D., Herzl A.,
RA Neumann S., Argitidou A., Vitale D., Ligouri R., Piravandi E.,
RA Massenet O., Quigley F., Glabaud G., Mendenhall A., Felder R.,
RA Schenati S., Hiller R., Schmidt W., Lechany A., Aubourg S.,
RA Chedori T., Cooke R., Berger C., Monfort A., Casauberta E.,
RA Gibbons T., Weber N., Vandenbol M., Baynes M., Terol J., Torres A.,
RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA Heijnen D., Schwarz S., Scholler P., Heber S., France P., Bleike C.,
RA Frishman D., Haase D., Lemcke K., Mews H.-W., Stocker S.,
RA Zaccaria P., Beyan M., Wilson R.K., de la Bastide M., Habermann K.,
RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
RA Sekhon M., Murray J., Shee P., Cordes M., Abu-Itiedeh U.,
RA Stoneking T., Kalicki J., Graves T., Hamon G., Edwards J.,
RA Lareelle P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA Mixx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
RA Nelson J., Spieth J., Ryan E., Andrews S., Gelsel C., Layman D.,
RA Du H., Ali J., Bergloff A., Jones K., Drone K., Cotton M., Joshu C.,
RA Antoniou B., Zidan M., Strong C., Sun H., Lamar B., Yordan C.,
RA Ma P., Zhong Y., Preston R., Vil D., Shekher M., Matero A., Shah R.,
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Tili S.,
RA Granat S., Shohdy N., Hasegawa A., Hamed A., Lodhi M., Johnson A.,
RA Chen E., Marra M., Martienssen R., McCombie W.R.,
RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
RT thaliana." ;
RL Nature 402:769-777 (1999) .
CC -I- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -I- TISSUE SPECIFICITY: STRONGLY UPREGULATED WITHIN DEVELOPING
CC FLOWERS, ESPECIALLY IN THE TAPETUM, THE DEVELOPING AND MATURE
CC POLLEN AND IN THE OVARIES.
CC -I- SIMILARITY: BELONGS TO THE PPP FAMILY OF PHOSPHATASES. PP-1
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on ways
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
CC or send an email to license@ebi.ac.uk).
CC -----
CC DR EMBL; Z46253; CAAB6339.1; -;
CC DR EMBL; U80921; AAC39460.1; -;
CC DR EMBL; AL096882; CAB51408.1; -;
CC DR EMBL; AL161531; CAB81225.1; -;
CC DR PIR; T13015; T13015.
CC DR HSSP; P08129; 1PJM.
CC DR InterPro; IPR004843; M-pyruvate.
CC DR InterPro; IPR006185; S/T phosphatase.
CC DR InterPro; IPR006186; T-phosphatase.
CC DR Pfam; PF00149; Metallophosphatase.
CC DR PRINTS; PR00114; STPHPTASE.
CC DR ProDom; PD000252; T-phosphatase.
CC DR SMART; SM00156; P2c; 1.
CC DR PROSITE; PS00125; SER_THR_PHOSPHATASE; 1.
CC KM Hydroxylase; Iron; Manganese; Multigene family.
CC FT METAL; 61 61 IRON (BY SIMILARITY).
CC FT METAL; 63 63 IRON (BY SIMILARITY).
CC FT METAL; 89 89 IRON AND MANGANESE (BY SIMILARITY).
CC FT ACT SITE; 121 121 MANGANESE (BY SIMILARITY).
CC FT METAL; 122 122 GENERAL ACID (BY SIMILARITY).
CC FT METAL; 170 170 MANGANESE (BY SIMILARITY).
CC FT METAL; 245 245 MANGANESE (BY SIMILARITY).
CC FT CONFLICT; 248 248 D -> V (IN REF. 3).
CC FT CONFLICT; 318 322 GSKGR -> V (IN REF. 3).
CC SO SEQUENCE; 322 AA; 36582 MW; B23128743E646205 CRC64;

Matches 13; Conservative 2; Mismatches 15; Indels 1; Gaps 1;
 Qy 8 DPSRK-KEMWENAGNKIYTMADKTIISKLMT 37
 |||:| || |||
 Db 205 DPREIQGNGENDRGVSYTFGADKVAEFLQT 235

Search completed: July 24, 2003, 11:54:26.
 Job time : 4.81552 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 11:47:53 ; Search time 8.9019 Seconds
(without alignments)
1159.539 Million cell updates/sec

Title: US-09-991-681-30

Perfect score: 215
Sequence: 1 SPKVEKKDPSRKKEWENAGNKIYTMADKTISKLMTEYK 40

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_protein:*
12: sp_virus:*
13: sp Vertebrate:*
14: sp Unclassified:*
15: sp_virus:*
16: sp_bacteria:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	215	100.0	589	4	Q8N4Y4
2	215	100.0	592	4	Q86CH9
3	215	100.0	1770	4	Q9ULH6
4	70	32.6	177	5	Q81995
5	67.5	31.4	193	5	Q819A5
6	67.5	31.4	743	5	Q81IU0
7	65	30.2	177	5	Q8AJE8
8	65	30.2	177	5	Q819A6
9	65	30.2	177	5	Q810M6
10	65	30.2	1062	10	Q8H8M1
11	64	29.8	621	2	Q59301
12	63.5	29.5	604	3	Q74491
13	61.5	28.6	181	5	Q819B1
14	61	28.4	3306	10	Q9FT44
15	59	27.4	177	5	Q81973
16	59	27.4	277	16	Q8U805

17	58	27.0	672	10	Q94IM7	Q94IM7 hordeum vul
18	57.5	26.7	921	16	Q929Z6	Q929Z6 listeria in
19	57	26.5	438	16	Q8PGZ8	Q8PGZ8 xanthomonas
20	57	26.5	683	11	Q8XV4	Q8XV4 mus musculu
21	56.5	26.3	201	16	Q8XBP3	Q8XBP3 escherichia
22	56.5	26.3	483	10	Q64951	Q64951 tagetes ere
23	56.5	26.3	574	10	Q64950	Q64950 tagetes ere
24	56	26.0	304	10	Q9SIS6	Q9SIS6 arabidopsis
25	56	26.0	322	16	Q8EWH2	Q8EWH2 mycoplasma
26	56	26.0	328	10	Q8S341	Q8S341 arabidopsis
27	56	26.0	328	10	Q8GWH8	Q8GWH8 arabidopsis
28	56	26.0	408	2	Q9LBP3	Q9LBP3 haemophilus
29	56	26.0	2162	5	Q8IEV1	Q8IEV1 plasmodium
30	56	26.0	2463	12	Q8BES6	Q8BES6 bovine kobu
31	55.5	25.8	181	5	Q819B2	Q819B2 plasmodium
32	55	25.6	177	10	Q24127	Q24127 nicotiana t
33	55	25.6	351	10	Q9MB29	Q9MB29 nicotiana t
34	55	25.6	351	10	Q9SMC1	Q9SMC1 nicotiana t
35	55	25.6	351	10	Q9SMB7	Q9SMB7 nicotiana t
36	55	25.6	654	10	Q9FL12	Q9FL12 arabidopsis
37	54.5	25.3	555	10	Q9AV96	Q9AV96 nicotiana t
38	54.5	25.3	673	10	Q94IM6	Q94IM6 hordeum vul
39	54.5	25.3	2120	5	Q81AK1	Q81AK1 plasmodium
40	54	25.1	175	5	Q81994	Q81994 plasmodium
41	54	25.1	732	2	Q8VP74	Q8VP74 cellvibrio
42	54	25.1	1048	10	Q04938	Q04938 oryza sativ
43	53.5	24.9	331	2	Q93EX6	Q93EX6 rhodococcus
44	53.5	24.9	1055	10	Q9LIV9	Q9LIV9 oryza sativ
45	53	24.7	174	5	Q81967	Q81967 plasmodium

ALIGNMENTS

RESULT 1	ID	Q8N4Y4	PRELIMINARY;	PRT;	589 AA.
AC	Q8N4Y4;				
DT	01-OCT-2002 (TREMBlrel. 22, Created)				
DT	01-OCT-2002 (TREMBlrel. 22, Last sequence update)				
DT	01-OCT-2002 (TREMBlrel. 22, Last annotation update)				
DE	Similar to KIAA1244 protein (Fragment).				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	(1)				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Colon;				
RA	Strausberg R.;				
RL	Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.				
DR	EMBL, BC033191; AAH33191.1; -.				
FT	NON_TER				
FT	NON_TER				
SQ	SEQUENCE 589 AA; 66086 MW; 8041EBA348DE5F7 CRC64;				
Query Match	100.0%; Score 215; DB 4; Length 589;				
Best Local Similarity	100.0%; Pred. No. 2.1e-19;				
Matches	40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Qy	1 SPKVEKKDPSRKKEWENAGNKIYTMADKTISKLMTEYK 40				
DB	403 SPKVEKKDPSRKKEWENAGNKIYTMADKTISKLMTEYK 442				
RESULT 2	ID	Q96CH9	PRELIMINARY;	PRT;	592 AA.
AC	Q96CH9;				
DT	01-DEC-2001 (TREMBlrel. 19, Created)				
DT	01-DEC-2001 (TREMBlrel. 19, Last sequence update)				
DT	01-OCT-2002 (TREMBlrel. 22, Last annotation update)				
DE	Hypothetical protein (Fragment).				
OS	Homo sapiens (Human).				

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strausberg R.;
RU Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC014227; AA014227.1; -.
KW Hypothetical protein.
FT NON TER
SQ SEQUENCE 592 AA; 66400 MW; F4A1E807B0DF47B5 CRC64;

Query Match
Best Local Similarity 100.0%; Score 215; DB 4; Length 592;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SPKVEKDPSSRKKEWENAGNKIYTMAADKTSKLTMTYK 40
Db 406 SPKVEKDPSSRKKEWENAGNKIYTMAADKTSKLTMTYK 445

RESULT 3
OYU0H6 PRELIMINARY; PRT; 1770 AA.
AC O9U0H6; Q96P46;
DT 01-MAR-2003 (TREMBLrel. 13, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE BIG3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RA Hong W.;
RT "K1A1244 as a novel distantly related member (BIG3) of the BIG1/Sec7p
RT subfamily of ARF GEFs.";
RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF413080; AA04174.1; -.
DR InterPro; IPR000904; Sec7.
DR SMART; SM00222; Sec7; 1.
KW Hypothetical protein.
SQ SEQUENCE 1770 AA; 195845 MW; SE996E36A6P92AB4 CRC64;

Query Match
Best Local Similarity 100.0%; Score 215; DB 4; Length 1770;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SPKVEKDPSSRKKEWENAGNKIYTMAADKTSKLTMTYK 40
Db 1584 SPKVEKDPSSRKKEWENAGNKIYTMAADKTSKLTMTYK 1623

RESULT 4
O81995 PRELIMINARY; PRT; 177 AA.
AC O81995;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Erythrocyte membrane protein 1 Gb174var8 (Fragment).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN (1)
RP SEQUENCE FROM N.A.
RA STRAIN=Gb174;
RA Khattab A., Kremsner P.G., Klinkert M.O.;
RT "Plasmodium falciparum placental isolates separated by time and space
RT express common surface antigen var genes of limited diversity.";
J. Infect. Dis. 0:0-0(2003).

```

```

DR EMBL; AF547116; AA06298.1; -.
FT NON TER
FT NON TER
SQ SEQUENCE 177 AA; 21163 MW; 09D26674C74AAA03 CRC64;

Query Match
Best Local Similarity 32.6%; Score 70; DB 5; Length 177;
Matches 14; Conservative 8; Mismatches 11; Indels 8; Gaps 1;

OY 6 KKDPSSRKKEWENAGNKIYTMAA-----DKTISKLTMTYK 38
Db 112 KKDPSSRKKEWENAGNKIYTMAA-----DKTISKLTMTYK 152

RESULT 5
O819A5 PRELIMINARY; PRT; 193 AA.
AC O819A5;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Erythrocyte membrane protein 1 Gb172var4 (Fragment).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN (1)
RP SEQUENCE FROM N.A.
RA STRAIN=Gb172;
RA Khattab A., Kremsner P.G., Klinkert M.O.;
RT "Plasmodium falciparum placental isolates separated by time and space
RT express common surface antigen var genes of limited diversity.";
J. Infect. Dis. 0:0-0(2003).
DR EMBL; AF547106; AA06288.1; -.
FT NON TER
FT NON TER
SQ SEQUENCE 193 AA; 22615 MW; 75DB2DCAFE9E34C42 CRC64;

Query Match
Best Local Similarity 31.4%; Score 67.5; DB 5; Length 193;
Matches 14; Conservative 9; Mismatches 15; Indels 13; Gaps 1;

OY 2 PKVEKDPSSRKKEWENAGNKIYTMA-----ADKTSKLTMTYK 39
Db 117 PTTKENTKQGEWENAGNKIYTMA-----ADKTSKLTMTYK 167

RESULT 6
O81U0 PRELIMINARY; PRT; 743 AA.
AC O81U0;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN (1)
RP SEQUENCE FROM N.A.
RA STRAIN=3D7;
RA MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Sub B., Peterson J., Angiuoli S.,
RA Pereira M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairclamb A.H., Fraunholz M.J., Rose D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Sudramanien G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Frazer C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum.";

```


Db 488 TPSSDSDLRCCOMNNNAKVATLEPDR 518

RESULT 11

059301 PRELIMINARY; PRT; 621 AA.

AC 059301; ID 059301; DT 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)

DR Endo-beta-1,4-xylanase precursor (EC 3.2.1.8).

OS Cellvibrio mixtus.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Cellvibrion.

NCBI_Taxid=39650;

CC NCBI_Taxid=39650;

CC (1)

CC SEQUENCE FROM N.A.

RC STRAIN=Mixtus;

RX MEDLINE=96077124; PubMed=7492333;

RA Millward-Sadler S.J., Davidson K., Hazlewood G.P., Black G.W., Gilbert H.J., Clarke J.H.;

RT "Novel cellulose-binding domains, NodB homologues and conserved modular architecture in xylanases from the aerobic soil bacteria Pseudomonas fluorescens subsp. cellulosa and Cellvibrio mixtus."

RT Biochem. J. 312:39-48(1995).

RL -1- SIMILARITY: BELONGS TO CELLULASE FAMILY F (FAMILY 10 OF GLYCOSYL HYDROLASES).

CC EMBL; Z48926; CAA88762.1; -.

DR HSSP; P14768; 1CLX.

DR InterPro; IPR005088; CBM_15.

DR InterPro; IPR001000; Glyco_hydro_10.

DR Pfam; PF03426; CBM_15; 1.

DR Pfam; PF00331; Glyco_hydro_10; 1.

DR SMART; SM00633; Glyco_10; 1.

DR SMART; SM00633; Glyco_10; 1.

DR PROSITE; PS00591; GLYCOSYL_HYDROL_F10; 1.

KW Glycosidase; Hydrolase; Signal; Xylan degradation.

FT SIGNAL 1 19 POTENTIAL.

SO SEQUENCE 621 AA; 64929 MW; 8305B959D2C3AB8 CRC64;

Query Match 29.8%; Score 64; DB 2; Length 621; Best Local Similarity 43.2%; Pred. No. 7.6; Matches 16; Conservative 5; Mismatches 16; Indels 0; Gaps 0;

QY 2 PVEKKDPSRKKEWENAGNKIYTMAADTKISKLMT 38

DB 175 PIVQIKGSGYPCBWCWAGNELETFAGEDATISCTVTE 211

RESULT 12

074491 PRELIMINARY; PRT; 604 AA.

AC 074491; P78908; ID 074491; DT 01-NOV-1998 (TREMBlrel. 08, Created)

DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)

DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)

DR Putative ATP-dependent RNA helicase C285.03.

OS Schizosaccharomyces pombe (fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes; Schizosaccharomycetales; Schizosaccharomycetaceae; Schizosaccharomycetes.

NCBI_Taxid=4896;

CC NCBI_Taxid=4896;

CC (1)

CC SEQUENCE FROM N.A.

RC STRAIN=C072;

RA Seeger K., Harris D., Lyne M., Rajandream M.A., Barrell B.G.;

RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE OF 115-604 FROM N.A.

RC STRAIN=PR745;

RX MEDLINE=98162722; PubMed=9501991;

Yoshioka S., Kato K., Nakai K., Okayama H., Nojima H.;

RT "Identification of open reading frames in Schizosaccharomyces pombe CDNA's."

RT DNA Res. 4:363-369(1997).

CC -1- SIMILARITY: TO OTHER "DEAD" BOX FAMILY HELICASES.

DR EMBL; AL031545; CAA20842.1; -.

DR EMBL; D89259; BAA13920.1; -.

DR GeneDB; SPombe; SPCC285.03; -.

DR InterPro; IPR001410; DEAD.

DR InterPro; IPR00629; DEAD_Box.

DR InterPro; IPR001650; Helicase_C.

DR Pfam; PF00270; DEAD; 1.

DR Pfam; PF00271; Helicase_C; 1.

DR SMART; SM00487; DEXDC; 1.

DR SMART; SM00490; HELIC; 1.

DR PROSITE; PS00039; DEAD_ATP_HELICASE; 1.

KW Hypothetical protein; ATP-binding; Helicase; RNA-binding.

FT NP BIND 180 187

FT SITE 292 295

FT CONFLICT 120 120

FT CONFLICT 123 123

FT CONFLICT 123 123

FT CONFLICT 401 401

FT CONFLICT 482 482

FT CONFLICT 502 504

FT CONFLICT 516 516

FT CONFLICT 527 527

FT CONFLICT 557 558

FT CONFLICT 578 579

FT CONFLICT 591 591

FT CONFLICT 596 596

SO SEQUENCE 604 AA; 68344 MW; 099005F2FD2CCD03 CRC64;

Query Match 29.5%; Score 63.5; DB 3; Length 604; Best Local Similarity 41.2%; Pred. No. 8.6; Matches 14; Conservative 5; Mismatches 14; Indels 1; Gaps 1;

QY 2 PVEKKDPSRKKEWENAGNKIYTMAADTKISKL 35

DB 13 PVEKKDPSRKKEWENAGNKIYTMAADTKISKL 45

RESULT 13

081981 PRELIMINARY; PRT; 181 AA.

AC 081981; ID 081981; DT 01-MAR-2003 (TREMBlrel. 23, Created)

DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)

DE Erythrocyte membrane protein 1 Gp170var4 (Fragment).

OS Plasmodium falciparum.

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

NCBI_Taxid=5833;

CC NCBI_Taxid=5833;

CC (1)

CC SEQUENCE FROM N.A.

RC STRAIN=Gb170;

RA Khattab A., Kremer P.G., Klinkert M.O.;

RT "Plasmodium falciparum placental isolates separated by time and space express common surface antigen var genes of limited diversity."

RT J. Infect. Dis. 0:0-0(2003).

DR EMBL; AF547100; AAN6282.1; -.

FT NON_TER 1 1

FT NON_TER 181 181

SO SEQUENCE 181 AA; 21277 MW; 64280214C7CA248B CRC64;

Query Match 28.6%; Score 61.5; DB 5; Length 181; Best Local Similarity 32.4%; Pred. No. 4.2; Matches 12; Conservative 8; Mismatches 14; Indels 3; Gaps 1;

QY 7 KDPSRKKEWENAGNKIY---TMAADTKISKLMTYK 40

DB 114 KDPSRKKEWENAGNKIY---TMAADTKISKLMTYK 150

RESULT 14

Q9FT44
ID Q9FT44 PRELIMINARY; PRT; 3306 AA.
AC Q9FT44;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE VPS13-like protein.
GN AT5G24740.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eustosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Murphy G., Ridley P., Hudson S., Bancroft I., Mewes H.W.,
RL Rudd S., Lemcke K., Mayer K.F.X.,
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL392145; CAC08248.1; --
DR IncePro; IPR001130; TAD_Dnae.
DR PROSITE; PS01090; TAD_2; 1.
SQ SEQUENCE 3306 AA; 368483 MW; 122B0D048E6C1E99 CRC64;

Query Match 28.4%; Score 61; DB 10; Length 3306;
Best Local Similarity 34.4%; Pred. No. 1.1e+02;
Matches 11; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

QY 1 SPKVEKDPSPRKEWENAGNKIYTMADKTI 32
DB 203 STSLSRKPPGWMKLMHYAQNVSLSIDSKI 234

RESULT 15

Q81973 PRELIMINARY; PRT; 177 AA.
ID Q81973
AC Q81973;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Erythrocyte membrane protein 1 Gb55var3 (Fragment).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Gb55;
RA Khattab A., Kremsner P.G., Klinkert M.O.;
RT "Plasmodium falciparum placental isolates separated by time and space
express common surface antigen var genes of limited diversity.",
RL J. Infect. Dis. 0:0-0(2003).
DR EMBL; AF547142; AAN86324.1; --
FT NON_TER 1
FT NON_TER 177
SQ SEQUENCE 177 AA; 20693 MW; E2E269567E24956C CRC64;

Query Match 27.4%; Score 59; DB 5; Length 177;
Best Local Similarity 31.8%; Pred. No. 8.7;
Matches 14; Conservative 9; Mismatches 15; Indels 6; Gaps 2;

QY 3 KVEKKDPSPRKEWENAGNKIY--TMADKTI--SKLMTPEYK 40
DB 108 KIGQSSIDQRKSWKEHGPEIWKGMCALEKTLDDKKKLNDRYK 151

Search completed: July 24, 2003, 11:57:06
Job time : 11.9019 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 11:58:19 ; Search time 8.91654 Seconds
(without alignments)
623.047 Million cell updates/sec

Title: US-09-991-681-29

Percent score: 35
Sequence: 1 SFGSSSTPTGTGFGKPTPEDDRSQSREHMGES 35

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1107863 seqs, 158726573 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

A_Geneseq_19Jun03:*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	100.0	35	AA019167	Human prostate-spe
2	35	100.0	35	AA050811	PS118 prostate mar
3	35	100.0	518	AA085472	PS118 prostate enco
4	35	100.0	518	AA019165	Human prostate-spe
5	35	100.0	518	AA050809	PS118 prostate mar
6	35	100.0	1807	ABG09728	Novel human diagno
7	35	100.0	1839	ABP64835	Human protein SEQ
8	35	100.0	1982	ABG09731	Novel human diagno
9	8	22.9	1031	AA035339	Enterococcus faeca

10	7	20.0	109	21	AA024569	Arabidopsis thalia
11	7	20.0	122	20	AA027570	Human secreted pro
12	7	20.0	160	21	AA024568	Arabidopsis thalia
13	7	20.0	166	20	AA026036	Secreted protein f
14	7	20.0	166	22	AA039003	Human secreted pro
15	7	20.0	166	23	AB055712	Human polypeptide
16	7	20.0	196	21	AA024567	Arabidopsis thalia
17	7	20.0	229	21	AA050805	Arabidopsis thalia
18	7	20.0	243	21	AA050804	Arabidopsis thalia
19	7	20.0	248	22	ABG16481	Novel human diagno
20	7	20.0	257	21	AA050803	Arabidopsis thalia
21	7	20.0	342	22	AA030416	C glutamicum prote
22	7	20.0	571	23	AB048070	Listeria monocytog
23	7	20.0	596	21	AA099408	Human PRO1342 (UNQ
24	7	20.0	596	22	AA029178	Human PRO polypept
25	7	20.0	596	22	AA087575	Human PRO1342. Ho
26	7	20.0	596	22	AA066157	Protein of the inv
27	7	20.0	596	23	ABG95900	Human secreted/tri
28	7	20.0	596	24	ABU71266	Human secreted pol
29	7	20.0	596	24	ABU71265	Human secreted pol
30	7	20.0	596	24	ABU72001	Novel human secret
31	7	20.0	596	24	ABU72158	Human PRO polypept
32	7	20.0	596	24	ABU65723	Human secreted/tri
33	7	20.0	596	24	ABU66056	Novel human secret
34	7	20.0	596	24	ABU67560	Human secreted/tri
35	7	20.0	596	24	ABU65418	Human PRO polypept
36	7	20.0	596	24	ABU58554	Human PRO polypept
37	7	20.0	596	24	ABU56090	Human secreted/tri
38	7	20.0	596	24	ABU57085	Human PRO polypept
39	7	20.0	596	24	ABU10664	Human secreted/tri
40	7	20.0	976	22	AB063981	Drosophila melanog
41	7	20.0	1310	22	AB063443	Drosophila melanog
42	7	20.0	2759	24	AA016418	Human nucleic acid
43	7	20.0	2781	21	AA057453	Human transcriptio
44	7	20.0	2907	21	AA057452	Human transcriptio
45	6	17.1	10	22	AA097469	Human complementar

ALIGNMENTS

RESULT 1	AA019167	standard; Peptide; 35 AA.
ID	AA019167	
XX	AC	AA019167;
XX	DT	27-NOV-2002 (first entry)
XX	DE	Human prostate-specific PS118 protein fragment #3.
XX	XX	
KM	KM	Human, prostate; prostate-specific sequence; prostate cancer; PS118;
XX	KM	EST; expressed sequence tag; cytosolic; gene therapy.
XX	OS	Homo sapiens.
PN	PN	US2002086316-A1.
XX	XX	
PD	PD	04-JUL-2002.
XX	XX	
PF	PF	26-NOV-2001; 2001US-0991681.
XX	XX	
PR	PR	23-APR-1998; 98US-0065383.
XX	XX	23-APR-1997; 97US-00842385.
XX	XX	
PA	PA	(BIL/) BILINGEL P A.
PA	PA	(COHE/) COHEN M.
PA	PA	(COLP/) COLPITTS T L.
PA	PA	(FRIE/) FRIEDMAN P N.
PA	PA	(GORD/) GORDAN J.
PA	PA	(GRAN/) GRANADOS E N.
PA	PA	(HODG/) HODGES S C.
PA	PA	(KLAS/) KLAS M R.

PA (KRAT/) KRATOCHVIL J D.
 PA (ROBE/) ROBERTS-RAPP L.
 PA (RUSS/) RUSSELL J C.
 PA (STRO/) STROUPE S D.
 XX
 PI Billingsel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J,
 PI Granados EN, Hodges SC, Klasse MR, Kratochvil JD, Roberts-Rapp L,
 PI Russell JC, Stroupe SD;
 XX
 DR WPI; 2002-665429/71.
 XX
 PT Novel PS118 polypeptide for detecting, diagnosing, staging, monitoring,
 PT prognosticating, preventing, treating, or determining predisposition of
 PT individual to diseases and conditions of prostate, e.g. prostate
 PT cancer -
 XX
 PS Claim 17; Page 44; 58pp; English.
 XX
 CC The present invention relates to a number of prostate-specific sequences
 CC derived from the human PS118 gene. These can be used in the detection,
 CC monitoring and treatment of prostate diseases, particularly prostate
 CC cancer. The present sequence is a PS118 protein fragment of the
 CC invention. The coding sequences of the invention were isolated from a
 CC prostate tissue expressed sequence tag (EST) library.
 XX
 SQ Sequence 35 AA;
 Query Match 100.0%; Score 35; DB 23; Length 35;
 Best Local Similarity 100.0%; Pred. No. 2.7e-29;
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SFOSESTPTGTGFGSGKETPSEDDRSQSRHEHMGES 35
 Db 1 SFOSESTPTGTGFGSGKETPSEDDRSQSRHEHMGES 35
 RESULT 2
 ID AAM50811 standard; Protein; 35 AA.
 XX
 AC AAM50811;
 XX
 DT 01-MAY-2002 (first entry)
 XX
 DE PS118 prostate marker immunogenic polypeptide.
 XX
 KW PS118; prostate; marker; prostate cancer; tumour; metastasis;
 KW benign prostatic hyperplasia; prostatic intraepithelial neoplasia;
 KW prostatitis; human; diagnosis; therapy; vaccine; immunogen.
 OS
 XX Homo sapiens.
 XX
 PN US2001055758-A1.
 XX
 PD 27-DEC-2001.
 XX
 PF 23-APR-1998; 98US-0065383.
 XX
 PR 23-APR-1997; 97US-0842385.
 XX
 PA (BILL/) BILLING-MEDEL P A.
 PA (COHE/) COHEN M.
 PA (CORL/) CORLETTIS T L.
 PA (FRIE/) FRIEDMAN P N.
 PA (GORD/) GORDON J.
 PA (GRAN/) GRANADOS E N.
 PA (HODG/) HODGES S C.
 PA (KLAS/) KLASSE M R.
 PA (KRAT/) KRATOCHVIL J D.
 PA (ROBE/) ROBERTS-RAPP L.
 XX
 PI Billings-Medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J,
 PI Granados EN, Hodges SC, Klasse MR, Kratochvil JD, Roberts-Rapp L,

PI Russell JC, Stroupe SD;
 XX
 DR WPI; 2002-187683/24.
 XX
 PT Detecting presence of target PS118 polynucleotide in test sample,
 PT useful for detecting, diagnosing, staging, monitoring, prognosticating,
 PT preventing or treating or determining predisposition to prostate
 PT disease -
 XX
 PS Claim 17; Page 44; 57pp; English.
 XX
 CC The present sequence is that of an immunogenic polypeptide
 CC comprising amino acids 283-317 of human prostate-specific PS118
 CC polypeptide (see AAM50809). A PS118 consensus sequence (see
 CC ABA91651) is found at least 12 times more often in prostate than
 CC in non-prostate tissue. PS118 polypeptides, polynucleotides,
 CC antibodies, agonists and inhibitors are useful for detecting,
 CC diagnosing, staging, monitoring, prognosticating, preventing and
 CC treating (including by genetic immunisation), or determining the
 CC predisposition of an individual to, diseases and conditions of the
 CC prostate, such as benign prostatic hyperplasia, prostatitis,
 CC prostatic intraepithelial neoplasia, prostate cancer, tumours and
 CC metastases. The PS118 polypeptides can be produced by expression
 CC of PS118 polynucleotides in transfected host cells, and
 CC immunogenic peptides are useful for raising PS118-specific
 CC antibodies of diagnostic use. The methods and reagents of the
 CC invention may provide an early means of detecting diseases of the
 CC prostate and may also provide new markers which can differentiate
 CC between the clinically important and unimportant prostate cancers
 CC without the use of surgery.
 XX
 SQ Sequence 35 AA;
 Query Match 100.0%; Score 35; DB 23; Length 35;
 Best Local Similarity 100.0%; Pred. No. 2.7e-29;
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SFOSESTPTGTGFGSGKETPSEDDRSQSRHEHMGES 35
 Db 1 SFOSESTPTGTGFGSGKETPSEDDRSQSRHEHMGES 35
 RESULT 3
 ID AAM85472 standard; Protein; 518 AA.
 XX
 AC AAM85472;
 XX
 DT 25-MAR-2003 (updated)
 DT 25-FEB-1999 (first entry)
 XX
 DE PS118 protein encoded by consensus sequence.
 XX
 KW EST clone; PS118; prostate tumour tissue; prostatic disease; cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO9848054-A1.
 XX
 PD 29-OCT-1998.
 XX
 PF 23-APR-1998; 98WO-US08239.
 XX
 PR 23-APR-1997; 97US-0842385.
 XX
 PA (ABBO) ABBOTT LAB.
 XX
 PI Billings-Medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J,
 PI Granados EN, Hodges SC, Klasse MR, Kratochvil JD, Robertstarrapp L,
 PI Russell JC, Stroupe SD;
 XX
 DR WPI; 1998-610000/51.
 DR N-PSDB; AAV82812.

XX New P118 nucleic acid and proteins - used for diagnosis and
PT treatment of prostatic disease, especially cancer, and also for drug
PT screening
XX
PS Claim 17; Page 93-94; 117pp; English.
XX
PS The present sequence is encoded by consensus P5118 sequence derived from
CC overlapping EST (expressed sequence tag) clones AAV82803-10. The clones
CC were identified from a cDNA library made from prostate tumour tissue.
CC Recombinant P5118 protein is used to detect P5118-specific antibodies,
CC to raise antibodies for detection of P5118 antigens, to screen for
CC specific binding agents (potential therapeutics), and to isolate specific
CC antibodies from serum. Detection of P5118 protein or nucleic acid, which
CC are prostate related, and altered or elevated in prostatic disease, is
CC used for detection, diagnosis, staging, monitoring and prognosis of
CC prostatic disease, particularly cancer, and to identify subjects at
CC risk.
CC (Updated on 25-MAR-2003 to correct PI field.)
CC
XX Sequence 518 AA;
SQ
Query Match 100.0%; Score 35; DB 19; Length 518;
Best Local Similarity 100.0%; Pred. No. 3.2e-28;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 SFOSESTPSTGSGKGTPEDDRSOSREHMGES 35
Db 283 SFOSESTPSTGSGKGTPEDDRSOSREHMGES 317
RESULT 4
AAO19165 standard; Protein: 518 AA.
XX
AC AAO19165;
XX
DT 27-NOV-2002 (first entry)
XX
DE Human prostate-specific P5118 protein fragment #1.
XX
KM Human; prostate; prostate-specific sequence; prostate cancer; P5118;
KM EST; expressed sequence tag; cytostatic; gene therapy.
XX
OS Homo sapiens.
XX
PN US2002086316-A1.
XX
PD 04-JUL-2002.
XX
PF 26-NOV-2001; 2001US-0991681.
XX
PR 23-APR-1998; 98US-0065383.
PR 23-APR-1997; 97US-0842385.
XX
PA (BILL/) BILLINGEL P A.
PA (COHE/) COHEN M.
PA (COLP/) COLPITTS T L.
PA (FRIE/) FRIEDMAN P N.
PA (GORD/) GORDAN J.
PA (GRAN/) GRANADOS E N.
PA (HODG/) HODGES S C.
PA (KLAS/) KLAS M R.
PA (KRAT/) KRATOCHVIL J D.
PA (ROBE/) ROBERTS-RAPP L.
PA (RUS/) RUSSELL J C.
PA (STRO/) STROUPE S D.
XX
XX Billingsel PA, Cohen M, Colpitts TL, Friedman PN, Gordan J;
PI Granados EN, Hodges SC, Kلاس MR, Kratochvil JD, Roberts-Rapp L;
PI Russell JC, Stroupe SD;
XX WPI; 2002-665429/71.
DR

XX Novel P5118 polypeptide for detecting, diagnosing, staging, monitoring,
PT prognosticating, preventing, treating, or determining predisposition of
PT individual to diseases and conditions of prostate, e.g. prostate
PT cancer -
XX
PS Claim 17; Page 42-43; 58pp; English.
XX
PS The present invention relates to a number of prostate-specific sequences
CC derived from the human P5118 gene. These can be used in the detection,
CC monitoring and treatment of prostate diseases, particularly prostate
CC cancer. The present sequence is a P5118 protein fragment of the
CC invention. The coding sequences of the invention were isolated from a
CC prostate tissue expressed sequence tag (EST) library.
XX
SQ Sequence 518 AA;
Query Match 100.0%; Score 35; DB 23; Length 518;
Best Local Similarity 100.0%; Pred. No. 3.2e-28;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 SFOSESTPSTGSGKGTPEDDRSOSREHMGES 35
Db 283 SFOSESTPSTGSGKGTPEDDRSOSREHMGES 317
RESULT 5
AAM50809 standard; Protein: 518 AA.
XX
AC AAM50809;
XX
DT 01-MAY-2002 (first entry)
XX
DE P5118 prostate marker partial sequence.
XX
KM P5118; prostate; marker; prostate cancer; tumour; metastasis;
KM benign prostatic hyperplasia; prostatic intraepithelial neoplasia;
KM prostatitis; human; diagnosis; therapy; vaccine.
XX
OS Homo sapiens.
XX
PN US2001055758-A1.
XX
PD 27-DEC-2001.
XX
PF 23-APR-1998; 98US-0065383.
XX
PR 23-APR-1997; 97US-0842385.
XX
PA (BILL/) BILLING-MEDEL P A.
PA (COHE/) COHEN M.
PA (COLP/) COLPITTS T L.
PA (FRIE/) FRIEDMAN P N.
PA (GORD/) GORDON J.
PA (GRAN/) GRANADOS E N.
PA (HODG/) HODGES S C.
PA (KLAS/) KLAS M R.
PA (KRAT/) KRATOCHVIL J D.
PA (ROBE/) ROBERTS-RAPP L.
XX
XX Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN, Gordan J;
PI Granados EN, Hodges SC, Kلاس MR, Kratochvil JD, Roberts-Rapp L;
PI Russell JC, Stroupe SD;
XX WPI; 2002-187683/24.
XX N-PSDB; ABA91651.
DR
XX Detecting presence of target P5118 polynucleotide in test sample,
PT useful for detecting, diagnosing, staging, monitoring, prognosticating,
PT preventing or treating or determining predisposition to prostate
PT disease -
XX

PS Claim 17; Page 42-43; 57bp; English.

XX The present sequence is that of a human prostate-specific PS118
CC polypeptide, as predicted from a partial consensus cDNA sequence
CC (see AB91651), and lacking the N-terminal region. The PS118
CC consensus sequence is found at least 12 times more often in
CC prostate than in non-prostate tissue. PS118 polypeptides,
CC including derivatives of the present sequence, polynucleotides,
CC antibodies, agonists and inhibitors are useful for detecting,
CC diagnosing, staging, monitoring, prognosticating, preventing and
CC treating (including by genetic immunisation), or determining the
CC predisposition of an individual to, diseases and conditions of the
CC prostate, such as benign prostatic hyperplasia, prostatitis, the
CC prostatic intraepithelial neoplasia, prostate cancer, tumours and
CC metastases. The PS118 polypeptides can be produced by expression
CC of PS118 polynucleotides in transfected host cells. The methods
CC and reagents of the invention may provide an early means of
CC detecting diseases of the prostate and may also provide new markers
CC which can differentiate between the clinically important and
CC unimportant prostate cancers without the use of surgery.

XX Sequence 518 AA;

XX SQ

Query Match 100.0%; Score 35; DB 23; Length 518;
Best Local Similarity 100.0%; Pred. No. 3.2e-26;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFQSSSTPSTGFGSGKETPSEDPRSQRHMGES 35
|||
Db 283 SFQSSSTPSTGFGSGKETPSEDPRSQRHMGES 317

RESULT 6
ABG09728
XX ABG09728 standard; Protein; 1807 AA.
XX AC
XX ABG09728;
XX DT 13-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #9719.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US08631.
XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX DR N-PSDB; AAS73915.
XX DR WPI: 2001-639362/73.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity
XX PS Claim 20; SEQ ID No 40087; 103bp; English.
XX CC The invention relates to isolated polynucleotide (I) and
XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1807 AA;

XX SQ

Query Match 100.0%; Score 35; DB 22; Length 1807;
Best Local Similarity 100.0%; Pred. No. 1e-27;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFQSSSTPSTGFGSGKETPSEDPRSQRHMGES 35
|||
Db 1572 SFQSSSTPSTGFGSGKETPSEDPRSQRHMGES 1606

RESULT 7
ABP64835
XX ABP64835 standard; Protein; 1839 AA.
XX AC
XX ABP64835;
XX DT 25-FEB-2003 (first entry)
XX DE Human protein SEQ ID 495.
XX KW Human; expressed sequence tag; EST;
XX KW haematopoietic disorder; central nervous system disease; viral infection;
XX KW peripheral nervous system disease; non-healing wound; infectious disease;
XX KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;
XX KW fungal infection; autoimmune disorder; coagulation disorder; noctropic;
XX KW antiallergic; antiinflammatory; immunosuppressive; neuroprotective;
XX KW cycostatic; haemostatic; virucide; antibacterial; fungicide;
XX KW immunostimulant; cerebroprotective.
XX OS Homo sapiens.
XX PN WO200259260-A2.
XX PD 01-AUG-2002.
XX PF 16-NOV-2001; 2001WO-US42950.
XX PR 17-NOV-2000; 2000US-0714936.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao Qa;
XX PI Ren F, Xue AJ, Yang Y, Weinman T, Drmanac RT;
XX DR N-PSDB; ABQ99421.
XX DR WPI: 2002-590824/63.
XX PT New isolated polynucleotide, useful in research, diagnostic or
XX PT therapeutic methods, e.g. preventing or treating disorders involving
XX PT aberrant protein expression or biological activity
XX PS Claim 20; SEQ ID 495; 394bp; English.

CC The present invention relates to novel human coding sequences
 CC (ABQ93268-ABQ93608) and proteins (ABP64682-ABP65022). The sequences are
 CC useful in therapeutic, diagnostic and research methods. The
 CC polynucleotides may be used in the field of molecular biology as
 CC hybridisation probes, primers for PCR, for chromosome and gene mapping,
 CC for the recombinant production of protein, or in generation of anti-sense
 CC DNA or RNA. The polynucleotides are useful in diagnostics as expressed
 CC sequence tags (ESTs) for identifying expressed genes or for physical
 CC mapping of the human genome. The proteins may be used as molecular weight
 CC markers, or as nutritional sources or supplements. The proteins may be
 CC used to maintain and expand cell population in a totipotential or
 CC pluripotential state useful for re-engineering damaged or diseased
 CC tissues, transplantation, manufacture of bio-pharmaceuticals or the
 CC development of bio-sensors. The polynucleotides and proteins are useful
 CC for preventing, treating or ameliorating disorders involving aberrant
 CC protein expression or biological activity, e.g. haematopoietic disorders,
 CC central/peripheral nervous system diseases, mechanical and traumatic
 CC disorders, non-healing wounds, immune deficiencies and disorders,
 CC infectious diseases caused by viral, bacterial or fungal infection,
 CC autoimmune disorders, allergic reactions and conditions, coagulation
 CC disorders, or cancer. The polynucleotide sequences of the invention were
 CC assembled from ESTs isolated mainly by sequencing by hybridisation, and
 CC in some cases, sequences obtained from one or more public databases.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 1839 AA;

Query Match 100.0%; Score 35; DB 23; Length 1839;
 Best Local Similarity 100.0%; Pred. No. 1e-27;

Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SFOSESTPSTGFGSGKETPSEDDRSOSRHHMGS 35
 Db 1604 SFOSESTPSTGFGSGKETPSEDDRSOSRHHMGS 1638

RESULT 8

ABG09731 ID ABG09731 standard; Protein; 1982 AA.

AC ABG09731;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #9722.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KM food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

DR N-PSDB; AAS73198.

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity

XX Claim 20; SEQ ID No 40090; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polynucleotide chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 1982 AA;

Query Match 100.0%; Score 35; DB 22; Length 1982;
 Best Local Similarity 100.0%; Pred. No. 1.1e-27;

Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SFOSESTPSTGFGSGKETPSEDDRSOSRHHMGS 35
 Db 1686 SFOSESTPSTGFGSGKETPSEDDRSOSRHHMGS 1720

RESULT 9

AAU35339 ID AAU35339 standard; Protein; 1031 AA.

AC AAU35339;

DT 14-FEB-2002 (first entry)

DE Enterococcus faecalis cellular proliferation protein #626.

KW Antisense; prokaryotic cellular proliferation protein;
 KM antibiotic; antibacterial; drug design.

OS Enterococcus faecalis.

PN WO200170955-A2.

PD 27-SEP-2001.

PF 21-MAR-2001; 2001WO-US09180.

PR 21-MAR-2000; 2000US-191078P.

PR 23-MAY-2000; 2000US-206848P.

PR 26-MAY-2000; 2000US-207727P.

PR 23-OCT-2000; 2000US-242578P.

PR 27-NOV-2000; 2000US-253625P.

PR 22-DEC-2000; 2000US-257931P.

PR 16-FEB-2001; 2001US-269308P.

PA (ELIT-) ELITRA PHARM INC.

PI Haselbeck R, Ohlsen KL, Zykkind JW, Wall D, Trawick JD, Carr GJ;

DR WPI; 2001-611495/70.

DR N-PSDB; AAS53198.

PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
PS Example 3; Seq ID No 10932; 511bp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1031 AA;
Query Match 22.9%; Score 8; DB 22; Length 1031;
Best Local Similarity 100.0%; Pred. No. 7.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 SSSSTPST 11
Db |||||
796 SSSSTPST 803
RESULT 10
AAG24569
ID AAG24569 standard; Protein; 109 AA.
XX
AC AAG24569;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 28293.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0125264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 99US-0131449.

PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134758.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-015086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147309.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149722.
 PR 23-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 25-AUG-1999; 99US-0149930.
 PR 26-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150864.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151348.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.

PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161927.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 20.0%; Score 7; DB 21; Length 109;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 STPSTGC 13
 Db 57 STPSTGC 63

RESULT 11
 AAY27570
 ID AAY27570 standard; Protein; 122 AA.

XX AAY27570;

XX 30-JUL-1999 (first entry)

DE Human secreted protein encoded by gene No. 4.

XX Human, secreted protein; fusion protein; gene therapy; protein therapy;
 KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
 KW developmental abnormality; foetal deficiency; blood; allergy; renal;
 KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
 KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.

OS Homo sapiens.

PN WO9924836-A1.

XX 20-MAY-1999.

PF 04-NOV-1998; 98WO-US23435.

PR 17-NOV-1997; 97US-0066100.
 PR 07-NOV-1997; 97US-0064900.
 PR 07-NOV-1997; 97US-0064908.
 PR 07-NOV-1997; 97US-0064911.
 PR 07-NOV-1997; 97US-0064912.
 PR 07-NOV-1997; 97US-0064983.
 PR 07-NOV-1997; 97US-0064984.
 PR 07-NOV-1997; 97US-0064985.
 PR 07-NOV-1997; 97US-0064987.
 PR 07-NOV-1997; 97US-0064988.
 PR 17-NOV-1997; 97US-0066090.
 PR 17-NOV-1997; 97US-0066094.
 PR 17-NOV-1997; 97US-0066095.
 PR 17-NOV-1997; 97US-0066089.

XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX

PI Carter KC, Edner R, Endress GA, Feng P, Janat F;
PI Kyaw H, Lafleur DW, Moore PA, Ni J, Olsen HS, Rosen CA;
PI Ruben SM, Shi Y, Soppet DR, Wei Y;
DR WPI: 1999-337740/28.
DR N-PSDB; AAX84936.
XX
PT New human secreted proteins and coding sequences useful for treating
PT disorders of the immune system and hyperproliferative disorders
PS Claim 11; Page 350-351; 507pp; English.
XX
CC This sequence represents a secreted human protein encoded by the gene
CC clone detailed in the descriptor line. The gene can be used to generate
CC fusion proteins by linking to the gene to a human immunoglobulin Fc
CC portion (e.g. AAX84924) for increasing the stability of the fused
CC protein as compared to the human protein only.
CC The invention relates to 125 novel genes and their fragments (nucleic
CC acid sequences: AAX84933-X85057; amino acid sequences AAX27567-Y27933)
CC which are useful for preventing, treating or ameliorating medical
CC conditions e.g. by protein or gene therapy. Also, pathological
CC conditions can be diagnosed by determining the amount of the new
CC polypeptides in a sample or by determining the presence of mutations in
CC the new polynucleotides. Specific uses are described for each of the 125
CC polynucleotides, based on which tissues they are most highly expressed in
CC (see AAX84933 for described uses).
XX
SQ Sequence 122 AA;
Query Match 20.0%; Score 7; DB 20; Length 122;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 15 SGETPS 21
Db 70 SGETPS 76
RESULT 12
AAG24568
ID AAG24568 standard; Protein; 160 AA.
XX
AC AAG24568;
XX
DT 17-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 28292.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
OS EP1033405-A2.
XX
FN
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134376.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136382.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142380.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148177.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151938.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.

PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 20.0%; Score 7; DB 21; Length 160;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 STPSTGG 13
DB 108 STPSTGG 114

RESULT 13
ID AAY26036
AC AAY26036;
DT 22-OCT-1999 (first entry)
XX
DE Secreted protein if87_1.
XX
KW Secreted protein; cytokine; cell proliferation; immune stimulation;
KW vaccine; immune suppression; haematopoiesis; tissue growth; activin;
KW inhibin; chemotaxis; chemokinesis; haemostasis; thrombolytic;
KW receptor; ligand; anti-inflammatory; cadherin; tumour; gene therapy.
XX
OS Homo sapiens.
XX
PN WO9928335-A1.
XX
PD 10-JUN-1999.
XX
PE 02-DEC-1998; 98WO-US25512.
XX
PR 30-NOV-1998; 98US-0203106.
PR 04-DEC-1997; 97US-0067454.
XX
PA (GENY) GENETICS INST INC.
XX
PI Agostino MJ, Clark HF, Collins-Racie LA, Evans C;
PI Pechtel K, Jacobs K, Lavallee ER, McCoy JM, Werberg D;
PI Steininger RJ, Treacy M, Wong GG;
DR WPI, 1999-385352/32.
DR N-PSDB; AAX80671.
XX
PT New polynucleotides encoding secreted human proteins
XX
PS Claim 25; Page 111; 124pp; English.
XX
CC The present sequence is a secreted protein if87_1, encoded by known clone
CC if87_1 (deposited as ATCC 98600) isolated from human adult uterus cDNA
CC library. Recombinant secreted proteins can be produced by

CC transforming host cells and culturing them under suitable conditions.
 CC The polynucleotide and protein are predicted to have biological
 CC activities which would make them suitable for treating, preventing or
 CC ameliorating medical conditions in humans and animals. Some predicted
 CC biological activities include cytokine and cell proliferation/
 CC differentiation activity, immune stimulating (e.g. as vaccines) or
 CC suppressing activity, hematopoiesis regulating activity, tissue growth
 CC activity, activin/inhibin activity, chemotactic/chemokinetic activity,
 CC hemostatic and thrombolytic activity, receptor/ligand activity, anti-
 CC inflammatory activity, cadherin/tumour invasion suppressor activity, and
 CC tumour inhibition activity. The polynucleotide encoding secreted
 CC protein can be used for gene therapy.

CC Sequence 186 AA;
 SQ

Query Match 20.0%; Score 7; DB 20; Length 186;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SESSSTPS 10
 |||||
 Db 123 SESSSTPS 129

RESULT 14
 AAU39003
 ID AAU39003 standard; Protein; 186 AA.
 XX
 XX AAU39003;
 AC
 XX
 XX 16-JAN-2002 (first entry)
 DT
 XX
 XX Human secreted protein if87_1.
 DE
 XX
 XX Human; secreted protein; antiinflammatory; immunosuppressive;
 KM neotropic; neuroprotective; antiarthritic; antimicrobial; vulnary;
 KM cytostatic; antidiabetic; virucide; antinfertility; anticonvulsant;
 KM vasotropic; antiparkinsonian; immunostimulant; dermatological;
 KM antirheumatic; antitumor; anticancer; osteopathic; tranquiliser;
 KM cerebroprotective; cytokine; cell proliferation; cell differentiation;
 KM immune deficiency; severe combined immunodeficiency; SCID; tumour;
 KM autoimmune disorder; multiple sclerosis; rheumatoid arthritis;
 KM graft-versus-host disease; myeloid deficiency; wound healing; ulcer;
 KM periodontal disease; osteoporosis; osteoarthritis; Alzheimer's disease;
 KM Parkinson's disease; Huntington's disease; infection; cardiac disease;
 KM stroke; sepsis; inflammatory bowel disease; contraceptive; immunogen;
 KM food supplement; vaccine.
 KM
 XX
 XX Homo sapiens.
 OS
 XX
 XX WO200175066-A2.
 PN
 XX
 XX 11-OCT-2001.
 PD
 XX
 XX 22-MAR-2001; 2001WO-US09369.
 PF
 XX
 XX 30-MAR-2000; 2000US-0539330.
 PR
 XX
 XX 04-DEC-2000; 2000US-0729674.
 XX
 XX (GENY) GENETICS INST INC.
 PA
 XX
 XX Jacobs K, McCoy JM, Lavallie E, Collins-racie LA, Evans C;
 PI Treacy M, Agostino MJ, Steininger RO, Spaulding V, Wong GG;
 PI Clark H, Fichtel K, Werberg D;
 XX
 XX WPI; 2001-639363/73.
 DR N-PSDB; AAS59221.
 DR
 XX
 XX Secreted human proteins, useful as vaccine for treating various
 PT diseases such as autoimmune disorders (e.g. multiple sclerosis), and
 PT nervous system disorders (e.g. stroke) -
 XX
 XX
 PS Disclosure; Page 476; 619pp; English.

XX The invention relates to novel human secreted proteins, the nucleic
 CC acids encoding them. The protein may exhibit cytokine, cell proliferation
 CC or cell differentiation activity or may induce production of other
 CC cytokines in certain cell populations and may exhibit immune stimulating
 CC or immune suppressing activity, which is useful for the treatment of
 CC various immune deficiencies and disorders e.g. severe combined
 CC immunodeficiency (SCID), autoimmune disorders e.g. multiple sclerosis,
 CC systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary
 CC inflammation. The proteins are also useful in the treatment of diseases
 CC and disorders including tissue, skin and organ transplantation and in
 CC graft-versus-host diseases (GVHD), in the induction of tumour immunity,
 CC myeloid or lymphoid cell deficiencies, wound healing and tissue repair.
 CC in the treatment of burns, incisions and ulcers; as well as in treatment
 CC of periodontal disease, osteoporosis or osteoarthritis, mediated by
 CC inflammatory processes, diseases of the peripheral nervous system,
 CC Alzheimer's, Parkinson's disease, Huntington's disease,
 CC amyotrophic lateral sclerosis, and Shy-Drager syndrome, infections,
 CC infarction of cardiac and central nervous system vessel e.g. stroke,
 CC sepsis, inflammatory bowel disease, ulcers, bone regeneration. The
 CC protein, having activin- or inhibin-related activities is useful as a
 CC contraceptive based on the ability of inhibin to decrease fertility in
 CC female mammals and decrease spermatogenesis in male mammals. The
 CC proteins and nucleic acids are also useful as food supplements. The
 CC present sequence represents a secreted protein of the invention.

CC Sequence 186 AA;
 SQ

Query Match 20.0%; Score 7; DB 22; Length 186;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SESSSTPS 10
 |||||
 Db 123 SESSSTPS 129

RESULT 15
 ABB5712
 ID ABB5712 standard; Protein; 186 AA.
 XX
 XX ABB5712;
 AC
 XX
 XX 14-FEB-2002 (first entry)
 DT
 XX
 XX Human polypeptide SEQ ID NO 30.
 DE
 XX
 XX Human; clone bd306-7; clone yb8-1; ATCC number 98599; gene therapy;
 KM immune disorder; bacterial infection; fungal infection; cancer; tumour;
 KM autoimmune disorder; systemic lupus erythematosus; wound; ulcer; inhibin;
 KM osteoporosis; osteoarthritis; nervous system disorder; neuropathy;
 KM Alzheimer's disease; Parkinson's disease; Huntington's disease; activin;
 KM haemophilia; cardiac infarction; stroke; sepsis; arthritis; vulnary;
 KM ischaemia-reperfusion injury; inflammatory bowel disease; chemotactic;
 KM Crohn's disease; cytostatic; anti-inflammatory; immunomodulatory;
 KM neuroprotective; haemostatic; thrombolytic; anti-inflammatory.
 KM
 XX
 XX Homo sapiens.
 OS
 XX
 XX US2001039335-A1.
 PN
 XX
 XX 08-NOV-2001.
 PD
 XX
 XX 04-DEC-2000; 2000US-0729674.
 PF
 XX
 XX 10-APR-1997; 97US-126425P.
 PR 04-DEC-1997; 97US-067454P.
 PR 20-DEC-1997; 97US-068379P.
 PR 02-01-1998; 98US-070346P.
 PR 07-JAN-1998; 98US-070643P.
 PR 08-JAN-1998; 98US-070755P.
 PR 13-JAN-1998; 98US-071304P.
 PR 22-JAN-1998; 98US-072134P.

PR 30-JAN-1998; 98US-073095P.
 PR 18-FEB-1998; 98US-075038P.
 PR 30-MAR-2000; 2000US-0539330.
 PR 23-NOV-1998; 98US-0197886C.

XX
 PA (JACO/) JACOBS K.
 PA (MCCO/) MCCOY J M.
 PA (LAVA/) LAVALLIE E R.
 PA (COLL/) COLLINS-RACIE L A.
 PA (EVAN/) EVANS C.
 PA (MERB/) MERBERG D.
 PA (TREA/) TREACY M.
 PA (AGOS/) AGOSTINO M J.
 PA (STEI/) STEININGER R J.
 PA (SPAU/) SPAULDING V.
 PA (WONG/) WONG G G.
 PA (CLAR/) CLARK H.
 PA (FECH/) FECHTEL K.

XX
 PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
 PI Merberg D, Treacy M, Agostino MJ, Steininger RJ, Spaulding V;
 PI Wong GG, Clark H, Fechtel K;
 DR N-PSDB; ABA90890.
 DR WPI; 2002-040725/05.

XX
 PT New secreted proteins and encoding polynucleotides, useful in gene
 PT therapies, particularly for preventing or treating autoimmune
 PT disorders, cancer, graft-versus-host disease, wound, osteoporosis,
 PT stroke or inflammations
 PS Disclosure; Page 197-198; 349pp; English.

XX
 CC The invention relates to isolated polynucleotides (ABA90876-ABA90968 and
 CC ABA90980) and encoded proteins (ABBS5698-ABBS5800), especially
 CC polynucleotides SEQ ID NO 1 (ABA90876) and SEQ ID NO 19 (ABA90885) and
 CC proteins SEQ ID NO 2 (ABBS5698) and SEQ ID NO 20 (ABBS5707) contained in
 CC clones bd306-7 and yb8-1 respectively and the clones bd306-7 and yb8-1
 CC are deposited with the American Type Culture Collection (ATCC) with
 CC accession number 98599. The polynucleotides and encoded polypeptides have
 CC cyostatic, anti-inflammatory, immunomodulator, vulnerary,
 CC neuroprotective, activin, inhibin, chemotactic, haemostatic, thrombolytic
 CC and anti-inflammatory activity and acting as cytokine modulators,
 CC haematopoiesis regulators, tissue growth modulators and/or cadherin
 CC suppressors. The polypeptides and polynucleotides are useful in gene
 CC therapies, particularly for preventing, treating or ameliorating any of
 CC the following diseases: immune deficiency and disorders; e.g. bacterial
 CC or fungal infections, autoimmune disorders, cancer, systemic lupus
 CC erythematosus or graft-versus-host disease; myeloid or lymphoid cell
 CC deficiencies; wound, burns, incisions and ulcers, osteoporosis or
 CC osteoarthritis; central and peripheral nervous system diseases and
 CC neuropathies, e.g. Alzheimer's, Parkinson's disease, Huntington's
 CC disease, amyotrophic lateral sclerosis or Sny-Drager syndrome;
 CC haemophilia, cardiac infarction or stroke; inflammations, shock, sepsis
 CC or systemic inflammatory response syndrome, ischaemia-reperfusion
 CC injury, endotoxin lethality, arthritis, inflammatory bowel disease or
 CC Crohn's disease; or tumours or cancers, pemphigus vulgaris or pemphigus
 CC foliaceus.

XX
 SQ Sequence 186 AA;

Query Match 20.0%; Score 7; DB 23; Length 186;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 SESSTPS 10
 |||||
 Db 123 SESSTPS 129

Search completed: July 24, 2003, 12:21:17
 Job time : 10.9165 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: July 24, 2003, 12:18:29 ; Search time 4.04832 Seconds
(without alignments)
365.801 Million cell updates/sec

Title: US-09-991-681-29

Perfect score: 35
Sequence: 1 SFQSESTPSTGFGSKETPSDDRSQSRHMGES 35

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310858 residues

Word size : 0

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	100.0	35	4	US-09-065-383-29
2	35	100.0	518	4	US-09-065-383-27
3	7	20.0	737	4	US-09-772-647-4
4	6	17.1	39	3	US-08-968-747-5
5	6	17.1	76	3	US-09-025-151-14
6	6	17.1	75	4	US-09-637-240-14
7	6	17.1	82	3	US-08-808-599A-37
8	6	17.1	94	2	US-08-828-008-2
9	6	17.1	105	4	US-09-252-991A-26885
10	6	17.1	118	4	US-09-252-991A-18728
11	6	17.1	123	4	US-09-732-210-1144
12	6	17.1	126	1	US-08-044-547-3
13	6	17.1	140	4	US-09-252-991A-28280
14	6	17.1	147	1	US-08-044-547-1
15	6	17.1	151	4	US-08-886-319A-58
16	6	17.1	153	4	US-09-886-319A-57
17	6	17.1	171	4	US-09-252-991A-30898
18	6	17.1	183	3	US-08-968-747-20
19	6	17.1	183	3	US-09-248-588-2
20	6	17.1	183	3	US-09-248-588-4
21	6	17.1	183	5	PCT-US96-10602-12
22	6	17.1	185	1	US-07-739-642-2
23	6	17.1	185	1	US-07-739-642-4
24	6	17.1	185	1	US-07-739-642-8
25	6	17.1	185	1	US-07-739-642-10
26	6	17.1	185	1	US-07-739-643-2
27	6	17.1	185	1	US-07-739-643-4

28	6	17.1	185	1	US-07-739-643-8	Sequence 8, Appli
29	6	17.1	185	1	US-07-739-643-10	Sequence 10, Appli
30	6	17.1	185	1	US-07-739-142-2	Sequence 2, Appli
31	6	17.1	185	1	US-07-739-142-4	Sequence 4, Appli
32	6	17.1	185	1	US-07-739-142-8	Sequence 8, Appli
33	6	17.1	185	1	US-07-739-142-10	Sequence 10, Appli
34	6	17.1	185	3	US-09-248-588-6	Sequence 6, Appli
35	6	17.1	185	4	US-09-851-120-6	Sequence 6, Appli
36	6	17.1	193	3	US-08-968-747-2	Sequence 2, Appli
37	6	17.1	194	3	US-08-968-747-18	Sequence 18, Appli
38	6	17.1	199	3	US-08-968-747-21	Sequence 21, Appli
39	6	17.1	211	4	US-09-252-991A-20225	Sequence 20225, A
40	6	17.1	211	6	5196194-13	Patent No. 5196194
41	6	17.1	212	3	US-08-968-747-3	Sequence 3, Appli
42	6	17.1	212	4	US-09-719-528A-4	Sequence 4, Appli
43	6	17.1	224	4	US-09-328-352-4411	Sequence 4411, Ap
44	6	17.1	270	4	US-09-389-913-59	Sequence 59, Appli
45	6	17.1	273	4	US-09-396-149-12	Sequence 12, Appli

ALIGNMENTS

RESULT 1
US-09-065-383-29
Sequence 29, Application US/09065383
Patent No. 6391543
GENERAL INFORMATION:
APPLICANT: BILTING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GORDON, JULIAN
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KLAAS, MICHAEL R.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: ROBERTS-RAPP, LISA
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
FOR DETECTING DISEASES OF THE PROSTATE
TITLE OF INVENTION: FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FaastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/065,383
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/842,385
FILING DATE: 23-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6084, US, P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 35 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6391543e
US-09-065-383-29

Query Match 100.0%; Score 35; DB 4; Length 35;
Best Local Similarity 100.0%; Pred. No. 6.4e-29;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SF0SESSTPSTGTFSGKETPSEDNRQSREHMGES 35
Db 1 SF0SESSTPSTGTFSGKETPSEDNRQSREHMGES 35

RESULT 2
US-09-065-383-27
Sequence 27, Application US/09065383
Patent No. 6391543

GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GORDON, JULIAN
APPLICANT: GRANDOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KLAS, MICHAEL R.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: ROBERTS-RAPP, LISA
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
TITLE OF INVENTION: FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0.
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/065,383
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/842,385
FILING DATE: 23-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6084.US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 518 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6391543e
US-09-065-383-27

Query Match 100.0%; Score 35; DB 4; Length 518;
Best Local Similarity 100.0%; Pred. No. 7.8e-28;

Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SF0SESSTPSTGTFSGKETPSEDNRQSREHMGES 35
Db 283 SF0SESSTPSTGTFSGKETPSEDNRQSREHMGES 317

RESULT 3
US-09-772-647-4
Sequence 4, Application US/09772647
Patent No. 6521815

GENERAL INFORMATION:
APPLICANT: Verna, Ajit K.
APPLICANT: Reddig, Peter J.
APPLICANT: Jansen, Aaron P.
TITLE OF INVENTION: Animal Model System for Squamous Cell Carcinoma
FILE REFERENCE: 960296.97613
CURRENT APPLICATION NUMBER: US/09/772,647
CURRENT FILING DATE: 2001-01-30
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 737
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: T7 tag and
OTHER INFORMATION: mouse protein kinase C epsilon coding sequence
US-09-772-647-4

Query Match 20.0%; Score 7; DB 4; Length 737;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 PSEDNR 26
Db 338 PSEDNR 344

RESULT 4
US-08-968-747-5

Sequence 5, Application US/08968747
Patent No. 6060595
GENERAL INFORMATION:
APPLICANT: Scaglioni et al.
TITLE OF INVENTION: INHIBITION OF VIRAL REPLICATION
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/968,747
FILING DATE: 03-SEP-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Frazer, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 08472/705001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:

LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-968-747-5

Query Match 17.1%; Score 6; DB 3; Length 39;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 RSQSR 30
Db 31 RSQSR 36

RESULT 5
US-09-025-151-14
Sequence 14, Application US/09025151
Patent No. 6187535
GENERAL INFORMATION:

APPLICANT: Legrain, Pierre
APPLICANT: Fromont, Micheline
TITLE OF INVENTION: FAST AND EXHAUSTIVE METHOD FOR SELECTING A PREY
TITLE OF INVENTION: POLYPEPTIDE INTERACTING WITH A BAIT POLYPEPTIDE OF
TITLE OF INVENTION: INTERACTORS POLYPEPTIDES
FILE REFERENCE: 03495-0164
CURRENT APPLICATION NUMBER: US/09/025.151
CURRENT FILING DATE: 1998-02-18
NUMBER OF SEQ ID NOS: 29
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 14
LENGTH: 76
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-09-025-151-14

Query Match 17.1%; Score 6; DB 3; Length 76;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 TPSED 24
Db 71 TPSED 76

RESULT 6
US-09-637-240-14
Sequence 14, Application US/09637240
Patent No. 6531284
GENERAL INFORMATION:

APPLICANT: Legrain, Pierre
APPLICANT: Fromont, Micheline
TITLE OF INVENTION: FAST AND EXHAUSTIVE METHOD FOR SELECTING A PREY
TITLE OF INVENTION: POLYPEPTIDE INTERACTING WITH A BAIT POLYPEPTIDE OF
TITLE OF INVENTION: INTERACTORS POLYPEPTIDES
FILE REFERENCE: 03495-0164
CURRENT APPLICATION NUMBER: US/09/637.240
CURRENT FILING DATE: 2000-06-14
PRIOR APPLICATION NUMBER: 09/025.151
PRIOR FILING DATE: 1998-02-18
NUMBER OF SEQ ID NOS: 29
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 14
LENGTH: 76
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-09-637-240-14

Query Match 17.1%; Score 6; DB 4; Length 76;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 TPSED 24
Db 71 TPSED 76

RESULT 7
US-08-808-599A-37

Sequence 37, Application US/0808599A
Patent No. 611089
GENERAL INFORMATION:
APPLICANT: Fukuda, Michiko N.
TITLE OF INVENTION: Trophinin, Trophinin-Assisting
TITLE OF INVENTION: Proteins and Methods to Inhibit Implantation
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell & Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/808.599A
FILING DATE: 28-FEB-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/317,522
FILING DATE: 04-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/439,818
FILING DATE: 12-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LA 2256
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 82 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-808-599A-37

Query Match 17.1%; Score 6; DB 3; Length 82;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 PSTGF 14
Db 65 PSTGF 70

RESULT 8
US-08-828-008-2

Sequence 2, Application US/08828008
Patent No. 5939316
GENERAL INFORMATION:
APPLICANT: Chen, Hong
APPLICANT: Friemert, Nelson, B.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
TITLE OF INVENTION: THE DIAGNOSIS AND TREATMENT OF NEUROPSYCHIATRIC
DISORDERS

NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036/2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/828,008
FILING DATE: 27-MAR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-071
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 94 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-828-008-2

Query Match 17.1%; Score 6; DB 2; Length 94;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 SSTPST 11
Db 13 SSTPST 18

RESULT 9
US-09-252-991A-26885
Sequence 26885, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 26885
LENGTH: 105
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26885

Query Match 17.1%; Score 6; DB 4; Length 105;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 SSTPST 11
Db 83 SSTPST 88

RESULT 10
US-09-252-991A-18728
Sequence 18728, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 18728
LENGTH: 118
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18728

Query Match 17.1%; Score 6; DB 4; Length 118;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 SSTPST 11
Db 45 SSTPST 50

RESULT 11
US-09-732-210-1144
Sequence 1144, Application US/09732210
Patent No. 6573361
GENERAL INFORMATION:
APPLICANT: Bunkers, Greg J.
APPLICANT: Liang, Jihong
APPLICANT: Mitmanek, Cindy A.
APPLICANT: Seale, Jeffrey W.
APPLICANT: Wu, Yonnie S.
TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use
FILE REFERENCE: 38-21(15036)B
CURRENT APPLICATION NUMBER: US/09/732,210
CURRENT FILING DATE: 2000-12-07
PRIOR APPLICATION NUMBER: US 60/169,513
PRIOR FILING DATE: 1999-12-07
PRIOR APPLICATION NUMBER: US 60/169,340
PRIOR FILING DATE: 1999-12-07
NUMBER OF SEQ ID NOS: 1753
SEQ ID NO 1144
LENGTH: 123
TYPE: PRT
ORGANISM: Clorella vulgaris
US-09-732-210-1144

Query Match 17.1%; Score 6; DB 4; Length 123;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 DRSOSR 29
Db 109 DRSOSR 114

RESULT 12
US-08-044-547-3
Sequence 3, Application US/08044547

Patent No. 5324715
GENERAL INFORMATION:
APPLICANT: Connolly, Thomas M.
APPLICANT: Keller, Paul M.
TITLE OF INVENTION: Protein for Inhibiting
COLLAGEN-STIMULATED PLATELET AGGREGATION
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000
CITY: Rahway
STATE: New Jersey
COUNTRY: US.
ZIP: 07065
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/044,547
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/594,917
FILING DATE: 09-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Parr, Richard S.
REGISTRATION NUMBER: 32,586
REFERENCE/DOCKET NUMBER: 18053
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-4958
TELEFAX: (908) 594-4720
TELEX: 138825
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 126 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-044-547-3

Query Match 17.1%; Score 6; DB 1; Length 126;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 TPSTGG 13
Db 24 TPSTGG 29

RESULT 13
US-09-252-991A-28280
Sequence 28280, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 28280
LENGTH: 140
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28280

Query Match 17.1%; Score 6; DB 4; Length 140;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 TPSTGG 12
Db 93 TPSTGG 98

RESULT 14
US-08-044-547-1
Sequence 1, Application US/08044547
Patent No. 5324715
GENERAL INFORMATION:
APPLICANT: Connolly, Thomas M.
APPLICANT: Keller, Paul M.
TITLE OF INVENTION: Protein for Inhibiting
COLLAGEN-STIMULATED PLATELET AGGREGATION
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000
CITY: Rahway
STATE: New Jersey
COUNTRY: US
ZIP: 07065
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/044,547
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/594,917
FILING DATE: 09-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Parr, Richard S.
REGISTRATION NUMBER: 32,586
REFERENCE/DOCKET NUMBER: 18053
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-4958
TELEFAX: (908) 594-4720
TELEX: 138825
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 147 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-044-547-1

Query Match 17.1%; Score 6; DB 1; Length 147;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 TPSTGG 13
Db 45 TPSTGG 50

RESULT 15
US-09-886-319A-58
Sequence 58, Application US/09886319A
Patent No. 6586185
GENERAL INFORMATION:
APPLICANT: Wolf, Eckard
APPLICANT: Werner, Sabine
APPLICANT: Halle, Jorn-Peter

```

; APPLICANT: Regenbogen, Johannes
; APPLICANT: Goppel, Andreas
; TITLE OF INVENTION: Use of Polypeptides or Nucleic Acids for
; TITLE OF INVENTION: the Diagnosis or Treatment of Skin Disorders and Wound
; TITLE OF INVENTION: Healing and for the Identification of Pharmacologically
; TITLE OF INVENTION: Active Substances
; FILE REFERENCE: 50125/014002
; CURRENT APPLICATION NUMBER: US/09/886,319A
; CURRENT FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: US 60/222,081
; PRIOR FILING DATE: 2000-08-01
; PRIOR APPLICATION NUMBER: DE 10030149.5
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: fastseq for Windows Version 4.0
; SEQ ID NO 58
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-886-319A-58

```

```

Query Match      17.1%; Score 6; DB 4; Length 151;
Best Local Similarity 100.0%; Pred.No. 50;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      11 TGGFSG 16      |||||
Db      81 TGGFSG 86

```

Search completed: July 24, 2003, 12:28:38
 Job time : 5.04832 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 12:26:04 ; Search time 5.94436 Seconds
(without alignments)
699.249 Million cell updates/sec

Title: US-09-991-681-29

Perfect score: 35
Sequence: 1 SFGSESTPTGTGFGSKETPSEDDRSQSRHMGES 35

Scoring table: OLIGO
Gapex 60.0, Gapext 60.0

Searched: 451899 seqs, 118759770 residues

Word size: 0

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

Published Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	22.9	1031	9	US-09-815-242-10932
2	7	20.0	121	11	US-09-974-879-143
3	7	20.0	122	11	US-09-305-736-142
4	7	20.0	186	9	US-09-729-674-30
5	7	20.0	342	10	US-09-738-626-4170
6	7	20.0	596	11	US-09-946-374-243
7	7	20.0	596	14	US-10-006-867-100
8	7	20.0	596	14	US-10-052-586-310
9	7	20.0	596	15	US-10-063-547-100
10	7	20.0	596	15	US-10-174-590-310
11	7	20.0	596	15	US-10-176-758-310
12	7	20.0	596	15	US-10-175-737-310
13	7	20.0	596	15	US-10-063-616-100
14	7	20.0	596	15	US-10-173-706-310
15	7	20.0	596	15	US-10-175-738-310

16	7	20.0	596	15	US-10-175-752-310	Sequence 310, App
17	7	20.0	596	15	US-10-176-482-310	Sequence 310, App
18	7	20.0	596	15	US-10-176-757-310	Sequence 310, App
19	7	20.0	596	15	US-10-176-913-310	Sequence 310, App
20	7	20.0	596	15	US-10-180-552-310	Sequence 310, App
21	7	20.0	596	15	US-10-180-557-310	Sequence 310, App
22	7	20.0	596	15	US-10-063-502-100	Sequence 100, App
23	7	20.0	596	15	US-10-173-700-310	Sequence 310, App
24	7	20.0	596	15	US-10-174-572-310	Sequence 310, App
25	7	20.0	596	15	US-10-174-579-310	Sequence 310, App
26	7	20.0	596	15	US-10-174-582-310	Sequence 310, App
27	7	20.0	596	15	US-10-174-588-310	Sequence 310, App
28	7	20.0	596	15	US-10-175-739-310	Sequence 310, App
29	7	20.0	596	15	US-10-175-740-310	Sequence 310, App
30	7	20.0	596	15	US-10-175-743-310	Sequence 310, App
31	7	20.0	596	15	US-10-176-488-310	Sequence 310, App
32	7	20.0	596	15	US-10-176-492-310	Sequence 310, App
33	7	20.0	596	15	US-10-176-747-310	Sequence 310, App
34	7	20.0	596	15	US-10-176-750-310	Sequence 310, App
35	7	20.0	596	15	US-10-176-985-310	Sequence 310, App
36	7	20.0	596	15	US-10-176-987-310	Sequence 310, App
37	7	20.0	596	15	US-10-176-992-310	Sequence 310, App
38	7	20.0	596	15	US-10-176-993-310	Sequence 310, App
39	7	20.0	596	15	US-10-184-658-310	Sequence 310, App
40	7	20.0	596	15	US-10-176-991-310	Sequence 310, App
41	7	20.0	596	15	US-10-173-695-310	Sequence 310, App
42	7	20.0	596	15	US-10-173-697-310	Sequence 310, App
43	7	20.0	596	15	US-10-173-705-310	Sequence 310, App
44	7	20.0	596	15	US-10-174-576-310	Sequence 310, App
45	7	20.0	596	15	US-10-174-585-310	Sequence 310, App

ALIGNMENTS

RESULT 1
US-09-815-242-10932
Sequence 10932, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haeselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Treawick, John D.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA 011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 1410
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10932
LENGTH: 1031
TYPE: PRT
ORGANISM: Enterococcus faecalis

US-09-815-242-10932

Query Match 22.9%; Score 8; DB 9; Length 1031;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SSSSTPST 11
|||||
DB 796 SSSSTPST 803

RESULT 2

US-09-974-879-143
; Sequence 143, Application US/09974879
; Publication No. US20030028003A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 125 Human Secreted Proteins
; FILE REFERENCE: P2020P2
; CURRENT APPLICATION NUMBER: US/09/974, 879
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/239, 893
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 09/818, 683
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 09/305, 736
; PRIOR FILING DATE: 1999-05-05
; PRIOR APPLICATION NUMBER: PCT/US98/23435
; PRIOR FILING DATE: 1998-11-04
; PRIOR APPLICATION NUMBER: US 60/064, 911
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064, 912
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064, 983
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064, 900
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064, 988
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064, 987
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064, 908
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064, 984
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/064, 985
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: US 60/066, 094
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: US 60/066, 100
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: US 60/066, 089
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: US 60/066, 095
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: US 60/066, 090
; PRIOR FILING DATE: 1997-11-17
; NUMBER OF SEQ ID NOS: 611
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 143
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-974-879-143

Query Match 20.0%; Score 7; DB 11; Length 121;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 SSKETPS 21
|||||
DB 70 SSKETPS 76

RESULT 3

US-09-305-736-142
; Sequence 142, Application US/09305736
; Publication No. US20030088078A1
; GENERAL INFORMATION:
; APPLICANT: Feng et al.
; TITLE OF INVENTION: 125 Human Secreted Proteins
; FILE REFERENCE: P2020P1
; CURRENT APPLICATION NUMBER: US/09/305, 736
; CURRENT FILING DATE: 1999-05-05
; EARLIER APPLICATION NUMBER: PCT/US98/23435
; EARLIER FILING DATE: 1998-11-04
; EARLIER APPLICATION NUMBER: 60/064, 911
; EARLIER FILING DATE: 1997-11-07
; EARLIER APPLICATION NUMBER: 60/064, 912
; EARLIER FILING DATE: 1997-11-07
; EARLIER APPLICATION NUMBER: 60/064, 983
; EARLIER FILING DATE: 1997-11-07
; EARLIER APPLICATION NUMBER: 60/064, 900
; EARLIER FILING DATE: 1997-11-07
; EARLIER APPLICATION NUMBER: 60/064, 988
; EARLIER FILING DATE: 1997-11-07
; EARLIER APPLICATION NUMBER: 60/064, 987
; EARLIER FILING DATE: 1997-11-07
; EARLIER APPLICATION NUMBER: 60/064, 908
; EARLIER FILING DATE: 1997-11-07
; EARLIER APPLICATION NUMBER: 60/064, 984
; EARLIER FILING DATE: 1997-11-07
; EARLIER APPLICATION NUMBER: 60/064, 985
; EARLIER FILING DATE: 1997-11-07
; EARLIER APPLICATION NUMBER: 60/066, 094
; EARLIER FILING DATE: 1997-11-17
; EARLIER APPLICATION NUMBER: 60/066, 100
; EARLIER FILING DATE: 1997-11-17
; EARLIER APPLICATION NUMBER: 60/066, 089
; EARLIER FILING DATE: 1997-11-17
; EARLIER APPLICATION NUMBER: 60/066, 095
; EARLIER FILING DATE: 1997-11-17
; EARLIER APPLICATION NUMBER: 60/066, 090
; EARLIER FILING DATE: 1997-11-17
; NUMBER OF SEQ ID NOS: 612
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 142
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (122)
; OTHER INFORMATION: Xaa equals stop translation
US-09-305-736-142

Query Match 20.0%; Score 7; DB 11; Length 122;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 SSKETPS 21
|||||
DB 70 SSKETPS 76

RESULT 4

US-09-729-674-30
; Sequence 30, Application US/09729674
; Patent No. US20010039335A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavallie, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David

```

; APPLICANT: Treacy, Maurice
; APPLICANT: Agostino, Michael J.
; APPLICANT: Steininger II, Robert J.
; APPLICANT: Spaulding, Vilki
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: 6055-64X
; CURRENT APPLICATION NUMBER: US/09/729,674
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: 09/539,330
; NUMBER OF SEQ ID NOS: 283
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 186
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-729-674-30

```

```

Query Match      20.0%; Score 7; DB 9; Length 186;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 4 SESSTPS 10
Db 123 SESSTPS 129

```

```

RESULT 5
US-09-738-626-4170
; Sequence 4170, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, MAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4170
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4170

```

```

Query Match      20.0%; Score 7; DB 10; Length 342;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 7 STPSTGG 13
Db 316 STPSTGG 322

```

```

RESULT 6
US-09-946-374-243
; Sequence 243, Application US/09946374
; Publication No. US20030073129A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Deenoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2830P1C1
; CURRENT APPLICATION NUMBER: US/09/946,374
; PRIOR FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: 60/098716
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098723
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098749
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098750
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098803
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/098821
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/098843
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/099536
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099596
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099598
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099602
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099642
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099741
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099754
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099763
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099792
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099808
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099812
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099815
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099816
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/100385
; PRIOR FILING DATE: 1998-09-15

```

PRIOR APPLICATION NUMBER: 60/100388
 PRIOR FILING DATE: 1998-09-15
 PRIOR APPLICATION NUMBER: 60/100390
 PRIOR FILING DATE: 1998-09-15
 PRIOR APPLICATION NUMBER: 60/100584
 PRIOR FILING DATE: 1998-09-16
 PRIOR APPLICATION NUMBER: 60/100627
 PRIOR FILING DATE: 1998-09-16
 PRIOR APPLICATION NUMBER: 60/100661
 PRIOR FILING DATE: 1998-09-16
 PRIOR APPLICATION NUMBER: 60/100662
 PRIOR FILING DATE: 1998-09-16
 PRIOR APPLICATION NUMBER: 60/100664
 PRIOR FILING DATE: 1998-09-16
 PRIOR APPLICATION NUMBER: 60/100683
 PRIOR FILING DATE: 1998-09-17
 PRIOR APPLICATION NUMBER: 60/100684
 PRIOR FILING DATE: 1998-09-17
 PRIOR APPLICATION NUMBER: 60/100710
 PRIOR FILING DATE: 1998-09-17
 PRIOR APPLICATION NUMBER: 60/100711
 PRIOR FILING DATE: 1998-09-17
 PRIOR APPLICATION NUMBER: 60/100848
 PRIOR FILING DATE: 1998-09-18
 PRIOR APPLICATION NUMBER: 60/100849
 PRIOR FILING DATE: 1998-09-18
 PRIOR APPLICATION NUMBER: 60/100919
 PRIOR FILING DATE: 1998-09-17
 PRIOR APPLICATION NUMBER: 60/100930
 PRIOR FILING DATE: 1998-09-17
 PRIOR APPLICATION NUMBER: 60/101014
 PRIOR FILING DATE: 1998-09-18
 PRIOR APPLICATION NUMBER: 60/101068
 PRIOR FILING DATE: 1998-09-18
 PRIOR APPLICATION NUMBER: 60/101071
 PRIOR FILING DATE: 1998-09-18
 PRIOR APPLICATION NUMBER: 60/101279
 PRIOR FILING DATE: 1998-09-22
 PRIOR APPLICATION NUMBER: 60/101471
 PRIOR FILING DATE: 1998-09-23
 PRIOR APPLICATION NUMBER: 60/101472
 PRIOR FILING DATE: 1998-09-23
 PRIOR APPLICATION NUMBER: 60/101474
 PRIOR FILING DATE: 1998-09-23
 PRIOR APPLICATION NUMBER: 60/101475
 PRIOR FILING DATE: 1998-09-23
 PRIOR APPLICATION NUMBER: 60/101476
 PRIOR FILING DATE: 1998-09-23
 PRIOR APPLICATION NUMBER: 60/101477
 PRIOR FILING DATE: 1998-09-23
 PRIOR APPLICATION NUMBER: 60/101479
 PRIOR FILING DATE: 1998-09-23
 PRIOR APPLICATION NUMBER: 60/101738
 PRIOR FILING DATE: 1998-09-24
 PRIOR APPLICATION NUMBER: 60/101741
 PRIOR FILING DATE: 1998-09-24
 PRIOR APPLICATION NUMBER: 60/101743
 PRIOR FILING DATE: 1998-09-24
 PRIOR APPLICATION NUMBER: 60/101915
 PRIOR FILING DATE: 1998-09-24
 PRIOR APPLICATION NUMBER: 60/101916
 PRIOR FILING DATE: 1998-09-24
 PRIOR APPLICATION NUMBER: 60/102207
 PRIOR FILING DATE: 1998-09-29
 PRIOR APPLICATION NUMBER: 60/102240
 PRIOR FILING DATE: 1998-09-29
 PRIOR APPLICATION NUMBER: 60/102307
 PRIOR FILING DATE: 1998-09-29
 PRIOR APPLICATION NUMBER: 60/102330
 PRIOR FILING DATE: 1998-09-29
 PRIOR APPLICATION NUMBER: 60/102331
 PRIOR FILING DATE: 1998-09-29
 PRIOR APPLICATION NUMBER: 60/102484

PRIOR FILING DATE: 1998-09-30
 PRIOR APPLICATION NUMBER: 60/102487
 PRIOR FILING DATE: 1998-09-30
 PRIOR APPLICATION NUMBER: 60/102570
 PRIOR FILING DATE: 1998-09-30
 PRIOR APPLICATION NUMBER: 60/102571
 PRIOR FILING DATE: 1998-09-30
 PRIOR APPLICATION NUMBER: 60/102684
 PRIOR FILING DATE: 1998-10-01
 PRIOR APPLICATION NUMBER: 60/102687
 PRIOR FILING DATE: 1998-10-01
 PRIOR APPLICATION NUMBER: 60/102965
 PRIOR FILING DATE: 1998-10-02
 PRIOR APPLICATION NUMBER: 60/103258
 PRIOR FILING DATE: 1998-10-06
 PRIOR APPLICATION NUMBER: 60/103314
 PRIOR FILING DATE: 1998-10-07
 PRIOR APPLICATION NUMBER: 60/103315
 PRIOR FILING DATE: 1998-10-07
 PRIOR APPLICATION NUMBER: 60/103328
 PRIOR FILING DATE: 1998-10-07
 PRIOR APPLICATION NUMBER: 60/103395
 PRIOR FILING DATE: 1998-10-07
 PRIOR APPLICATION NUMBER: 60/103396
 PRIOR FILING DATE: 1998-10-07
 PRIOR APPLICATION NUMBER: 60/103401
 PRIOR FILING DATE: 1998-10-07
 PRIOR APPLICATION NUMBER: 60/103449
 PRIOR FILING DATE: 1998-10-06
 PRIOR APPLICATION NUMBER: 60/103633
 PRIOR FILING DATE: 1998-10-08
 PRIOR APPLICATION NUMBER: 60/103678
 PRIOR FILING DATE: 1998-10-08
 PRIOR APPLICATION NUMBER: 60/103679
 PRIOR FILING DATE: 1998-10-08
 PRIOR APPLICATION NUMBER: 60/103711
 PRIOR FILING DATE: 1998-10-08
 PRIOR APPLICATION NUMBER: 60/104257
 PRIOR FILING DATE: 1998-10-14
 PRIOR APPLICATION NUMBER: 60/104987
 PRIOR FILING DATE: 1998-10-20
 PRIOR APPLICATION NUMBER: 60/105000
 PRIOR FILING DATE: 1998-10-20
 PRIOR APPLICATION NUMBER: 60/105002
 PRIOR FILING DATE: 1998-10-20
 PRIOR APPLICATION NUMBER: 60/105104
 PRIOR FILING DATE: 1998-10-21
 PRIOR APPLICATION NUMBER: 60/105169
 PRIOR FILING DATE: 1998-10-22
 PRIOR APPLICATION NUMBER: 60/105266
 PRIOR FILING DATE: 1998-10-22
 PRIOR APPLICATION NUMBER: 60/105693
 PRIOR FILING DATE: 1998-10-26
 PRIOR APPLICATION NUMBER: 60/105694
 PRIOR FILING DATE: 1998-10-26
 PRIOR APPLICATION NUMBER: 60/105807

Query Match 20.0%; Score 7; DB 11; Length 596;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 SESSTPS 10
 Db 123 SESSTPS 129
 RESULT 7
 US-10-006-867-100
 ; Sequence 100; Application US/10006867
 ; Publication No. US20020119130A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Eaton, Dan L.
 ; APPLICANT: Filvaroff, Ellen

APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Watanabe, Colin K.
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3230R1C1
CURRENT APPLICATION NUMBER: US/10/006,867
CURRENT FILING DATE: 2001-12-06
PRIOR APPLICATION NUMBER: 60/063435
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/064215
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088740
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088811
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088825
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088863
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090688
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091628
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/096012
PRIOR FILING DATE: 1998-08-10
PRIOR APPLICATION NUMBER: 60/096757
PRIOR FILING DATE: 1998-08-17
PRIOR APPLICATION NUMBER: 60/096949
PRIOR FILING DATE: 1998-08-18
PRIOR APPLICATION NUMBER: 60/096959
PRIOR FILING DATE: 1998-08-18
PRIOR APPLICATION NUMBER: 60/097954
PRIOR FILING DATE: 1998-08-26
PRIOR APPLICATION NUMBER: 60/097971

PRIOR FILING DATE: 1998-08-26
PRIOR APPLICATION NUMBER: 60/097979
PRIOR FILING DATE: 1998-08-26
PRIOR APPLICATION NUMBER: 60/098749
PRIOR FILING DATE: 1998-09-01
PRIOR APPLICATION NUMBER: 60/099741
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099763
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099792
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099812
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099815
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/100627
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 60/100662
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 60/100683
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/100684
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/100930
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/101279
PRIOR FILING DATE: 1998-09-22
PRIOR APPLICATION NUMBER: 60/101475
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101738
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101743
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101916
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/102570
PRIOR FILING DATE: 1998-09-30
PRIOR APPLICATION NUMBER: 60/103449
PRIOR FILING DATE: 1998-10-06
PRIOR APPLICATION NUMBER: 60/103678
PRIOR FILING DATE: 1998-10-08
PRIOR APPLICATION NUMBER: 60/103679
PRIOR FILING DATE: 1998-10-08
PRIOR APPLICATION NUMBER: 60/103711
PRIOR FILING DATE: 1998-10-08
PRIOR APPLICATION NUMBER: 60/105000
PRIOR FILING DATE: 1998-10-20
PRIOR APPLICATION NUMBER: 60/105002
PRIOR FILING DATE: 1998-10-20
PRIOR APPLICATION NUMBER: 60/105881
PRIOR FILING DATE: 1998-10-27
PRIOR APPLICATION NUMBER: 60/106030
PRIOR FILING DATE: 1998-10-28
PRIOR APPLICATION NUMBER: 60/106464
PRIOR FILING DATE: 1998-10-30
PRIOR APPLICATION NUMBER: 60/106856
PRIOR FILING DATE: 1998-11-03
PRIOR APPLICATION NUMBER: 60/108807
PRIOR FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: 60/112419
PRIOR FILING DATE: 1998-12-15
PRIOR APPLICATION NUMBER: 60/112422
PRIOR FILING DATE: 1998-12-15
PRIOR APPLICATION NUMBER: 60/112853
PRIOR FILING DATE: 1998-12-16
PRIOR APPLICATION NUMBER: 60/113011
PRIOR FILING DATE: 1998-12-16
PRIOR APPLICATION NUMBER: 60/112854
PRIOR FILING DATE: 1998-12-16
PRIOR APPLICATION NUMBER: 60/113300
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/113408
PRIOR FILING DATE: 1998-12-22

PRIOR APPLICATION NUMBER: 60/113430
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/113621
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/114223
PRIOR FILING DATE: 1998-12-30
PRIOR APPLICATION NUMBER: 60/115614
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/116527
PRIOR FILING DATE: 1999-01-20
PRIOR APPLICATION NUMBER: 60/116843
PRIOR FILING DATE: 1999-01-22
PRIOR APPLICATION NUMBER: 60/119285
PRIOR FILING DATE: 1999-02-09
PRIOR APPLICATION NUMBER: 60/119287
PRIOR FILING DATE: 1999-02-09
PRIOR APPLICATION NUMBER: 60/119525
PRIOR FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: 60/119549
PRIOR FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: 60/120014
PRIOR FILING DATE: 1999-02-11
PRIOR APPLICATION NUMBER: 60/129122
PRIOR FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: 60/129674
PRIOR FILING DATE: 1999-04-16
PRIOR APPLICATION NUMBER: 60/131291
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/138387
PRIOR FILING DATE: 1999-06-09
PRIOR APPLICATION NUMBER: 60/144791
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/169495
PRIOR FILING DATE: 1999-12-07
PRIOR APPLICATION NUMBER: 60/175481
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: 60/191007
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/199397
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/380139
PRIOR FILING DATE: 1998-08-25
PRIOR APPLICATION NUMBER: 09/311832
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 09/380137
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/380138
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/380142

Query Match 20.0%; Score 7; DB 14; Length 596;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0;
Qy 4 SESSTPS 10
DB 123 SESSTPS 129

RESULT 8
US-10-052-586-310
Sequence 310, Application US/10052586
Publication No. US20020127584A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.

APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P343OR1C1
CURRENT APPLICATION NUMBER: US/10/052,586
CURRENT FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059266
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063120
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063121
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063486
PRIOR FILING DATE: 1997-10-21
PRIOR APPLICATION NUMBER: 60/063540
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063541
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063544
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063564
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063734
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063870
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/064103
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066120
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066466
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066772
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/069335
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069425
PRIOR FILING DATE: 1997-12-12
PRIOR APPLICATION NUMBER: 60/069870
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: 60/068017
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08

;; PRIOR APPLICATION NUMBER: 60/081070
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081195
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081838
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/082568
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082569
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082704
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082797
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/083322
;; PRIOR FILING DATE: 1998-04-28
;; PRIOR APPLICATION NUMBER: 60/083495
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083496
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083499
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083559
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083366
;; PRIOR FILING DATE: 1998-05-05
;; PRIOR APPLICATION NUMBER: 60/084414
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084639
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084640
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084643
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/085573
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085580
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085582
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085700
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/086023
;; PRIOR FILING DATE: 1998-05-16
;; PRIOR APPLICATION NUMBER: 60/086392
;; PRIOR FILING DATE: 1998-05-22
;; PRIOR APPLICATION NUMBER: 60/086486
;; PRIOR FILING DATE: 1998-05-22
;; PRIOR APPLICATION NUMBER: 60/087098
;; PRIOR FILING DATE: 1998-05-28
;; PRIOR APPLICATION NUMBER: 60/087208
;; PRIOR FILING DATE: 1998-05-28
;; PRIOR APPLICATION NUMBER: 60/087609
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087759
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087827
;; PRIOR FILING DATE: 1998-06-03
;; PRIOR APPLICATION NUMBER: 60/088025
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088028
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088029
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088033
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088167
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088202
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088212

;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088217
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088326
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088655
;; PRIOR FILING DATE: 1998-06-09
;; PRIOR APPLICATION NUMBER: 60/088722
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088738
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088740
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088811
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088824
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088825
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088826
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088861
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088863
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088876
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/089090
;; PRIOR FILING DATE: 1998-06-12
;; PRIOR APPLICATION NUMBER: 60/089105
;; PRIOR FILING DATE: 1998-06-12
;; PRIOR APPLICATION NUMBER: 60/089512
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089514
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089538
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089598
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089653
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089908

Query Match 20.0%; Score 7; DB 14; Length 596;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SESSTPS 10
Db 123 SESSTPS 129

RESULT 9
US-10-063-547-100
; Sequence 100, Application US/10063547
; Publication No. US20020182638A1
; GENERAL INFORMATION:
; APPLICANT: Baton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,547
; PRIOR FILING DATE: 2002-05-02
; Prior Application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 170

SEQ ID NO 100
LENGTH: 596
TYPE: PRT
ORGANISM: Homo Sapien
US-10-063-547-100

Query Match 20.0%; Score 7; DB 15; Length 596;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SESSSTPS 10
|||||
Db 123 SESSSTPS 129

RESULT 10
US-10-174-590-310

; Sequence 310, Application US/10174590
; Publication No. US20030008352A1

GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C42
; CURRENT APPLICATION NUMBER: US/10/174,590
; CURRENT FILING DATE: 2002-06-18
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-590-310

Query Match 20.0%; Score 7; DB 15; Length 596;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SESSSTPS 10
|||||
Db 123 SESSSTPS 129

RESULT 11
US-10-176-758-310

; Sequence 310, Application US/10176758
; Publication No. US20030008353A1

GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C104
; CURRENT APPLICATION NUMBER: US/10/176,758

; CURRENT FILING DATE: 2002-06-21
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-758-310

Query Match 20.0%; Score 7; DB 15; Length 596;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SESSSTPS 10
|||||
Db 123 SESSSTPS 129

RESULT 12
US-10-175-737-310

; Sequence 310, Application US/10175737
; Publication No. US20030013153A1

GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C50
; CURRENT APPLICATION NUMBER: US/10/175,737
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-737-310

Query Match 20.0%; Score 7; DB 15; Length 596;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SESSSTPS 10
|||||
Db 123 SESSSTPS 129

RESULT 13
US-10-063-616-100

; Sequence 100, Application US/10063616
; Publication No. US20030013855A1

GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Getlitsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3230R1C1

```

; CURRENT APPLICATION NUMBER: US/10/063,616
; CURRENT FILING DATE: 2002-05-03
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 100
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-063-616-100

```

```

Query Match      20.0%; Score 7; DB 15; Length 596;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      4 SESSTPS 10
        |||||
Db      123 SESSTPS 129

```

```

RESULT 14
US-10-173-706-310
; Sequence 310, Application US/10173706
; Publication No. US20030022293a1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C7
; CURRENT APPLICATION NUMBER: US/10/173,706
; CURRENT FILING DATE: 2002-06-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-173-706-310

```

```

Query Match      20.0%; Score 7; DB 15; Length 596;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      4 SESSTPS 10
        |||||
Db      123 SESSTPS 129

```

```

RESULT 15
US-10-175-738-310
; Sequence 310, Application US/10175738
; Publication No. US20030022294a1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin

```

```

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C45
; CURRENT APPLICATION NUMBER: US/10/175,738
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-738-310

```

```

Query Match      20.0%; Score 7; DB 15; Length 596;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      4 SESSTPS 10
        |||||
Db      123 SESSTPS 129

```

Search completed: July 24, 2003, 12:48:28
Job time : 6.94436 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: July 24, 2003, 12:21:29 ; Search time 52.7306 Seconds
(without alignment)
577.701 Million cell updates/sec

Title: US-09-991-681-29

Perfect score: 35
Sequence: 1 SFQSESTPTSGFGSGKETPSEDSDRSQSRHMGES 35

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 5580241 seqs, 870357830 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5580241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Pending Patents_Aa Main:*

1: /cgn2_6/ptodata/1/paa/PCRTUS_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US099A_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US099B_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
28: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
29: /cgn2_6/ptodata/1/paa/US103_COMB.pep.*
30: /cgn2_6/ptodata/1/paa/US104_COMB.pep.*
31: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	100.0	35	US-08-842-385-8	Sequence 8, Appl1
2	35	100.0	35	US-09-991-681-29	Sequence 29, Appl1

3	35	100.0	467	12	US-08-842-385-6	Sequence 6, Appl1
4	35 <th>100.0</th> <td>518</td> <td>25<td>US-09-991-681-27</td><td>Sequence 27, Appl1</td></td>	100.0	518	25 <td>US-09-991-681-27</td> <td>Sequence 27, Appl1</td>	US-09-991-681-27	Sequence 27, Appl1
5	35 <th>100.0</th> <td>1770</td> <td>1<td>PCT-US03-01943-44</td><td>Sequence 44, Appl1</td></td>	100.0	1770	1 <td>PCT-US03-01943-44</td> <td>Sequence 44, Appl1</td>	PCT-US03-01943-44	Sequence 44, Appl1
6	35 <th>100.0</th> <td>1770</td> <td>27<td>US-10-144-198-44</td><td>Sequence 44, Appl1</td></td>	100.0	1770	27 <td>US-10-144-198-44</td> <td>Sequence 44, Appl1</td>	US-10-144-198-44	Sequence 44, Appl1
7	35 <th>100.0</th> <td>1807</td> <td>1<td>PCT-US01-08631-40087</td><td>Sequence 40087, A</td></td>	100.0	1807	1 <td>PCT-US01-08631-40087</td> <td>Sequence 40087, A</td>	PCT-US01-08631-40087	Sequence 40087, A
8	35 <th>100.0</th> <td>1839</td> <td>1<td>PCT-US01-42950-495</td><td>Sequence 495, App</td></td>	100.0	1839	1 <td>PCT-US01-42950-495</td> <td>Sequence 495, App</td>	PCT-US01-42950-495	Sequence 495, App
9	35 <th>100.0</th> <td>1839</td> <td>30<td>US-10-416-993-495</td><td>Sequence 32, Appl1</td></td>	100.0	1839	30 <td>US-10-416-993-495</td> <td>Sequence 32, Appl1</td>	US-10-416-993-495	Sequence 32, Appl1
10	35 <th>100.0</th> <td>1872</td> <td>1<td>PCT-US03-04508-32</td><td>Sequence 32, Appl1</td></td>	100.0	1872	1 <td>PCT-US03-04508-32</td> <td>Sequence 32, Appl1</td>	PCT-US03-04508-32	Sequence 32, Appl1
11	35 <th>100.0</th> <td>1982</td> <td>1<td>PCT-US01-06631-40090</td><td>Sequence 40090, A</td></td>	100.0	1982	1 <td>PCT-US01-06631-40090</td> <td>Sequence 40090, A</td>	PCT-US01-06631-40090	Sequence 40090, A
12	35 <th>100.0</th> <td>2221</td> <td>1<td>PCT-US03-01943-30</td><td>Sequence 30, Appl1</td></td>	100.0	2221	1 <td>PCT-US03-01943-30</td> <td>Sequence 30, Appl1</td>	PCT-US03-01943-30	Sequence 30, Appl1
13	35 <th>100.0</th> <td>2221</td> <td>27<td>US-10-144-198-30</td><td>Sequence 30, Appl1</td></td>	100.0	2221	27 <td>US-10-144-198-30</td> <td>Sequence 30, Appl1</td>	US-10-144-198-30	Sequence 30, Appl1
14	20 <th>57.1</th> <td>141</td> <td>28<td>US-10-221-279-7783</td><td>Sequence 7783, Ap</td></td>	57.1	141	28 <td>US-10-221-279-7783</td> <td>Sequence 7783, Ap</td>	US-10-221-279-7783	Sequence 7783, Ap
15	8 <th>22.9</th> <td>270</td> <td>15<td>US-09-134-000-3553</td><td>Sequence 3553, Ap</td></td>	22.9	270	15 <td>US-09-134-000-3553</td> <td>Sequence 3553, Ap</td>	US-09-134-000-3553	Sequence 3553, Ap
16	8 <th>22.9</th> <td>270</td> <td>15<td>US-09-134-000-3553</td><td>Sequence 3553, Ap</td></td>	22.9	270	15 <td>US-09-134-000-3553</td> <td>Sequence 3553, Ap</td>	US-09-134-000-3553	Sequence 3553, Ap
17	8 <th>22.9</th> <td>270</td> <td>30<td>US-10-434-665-3553</td><td>Sequence 3553, Ap</td></td>	22.9	270	30 <td>US-10-434-665-3553</td> <td>Sequence 3553, Ap</td>	US-10-434-665-3553	Sequence 3553, Ap
18	8 <th>22.9</th> <td>1031</td> <td>1<td>PCT-US02-03987-10932</td><td>Sequence 10932, A</td></td>	22.9	1031	1 <td>PCT-US02-03987-10932</td> <td>Sequence 10932, A</td>	PCT-US02-03987-10932	Sequence 10932, A
19	8 <th>22.9</th> <td>1031</td> <td>23<td>US-09-815-242-10932</td><td>Sequence 10932, A</td></td>	22.9	1031	23 <td>US-09-815-242-10932</td> <td>Sequence 10932, A</td>	US-09-815-242-10932	Sequence 10932, A
20	8 <th>22.9</th> <td>1031</td> <td>26<td>US-10-072-851-10932</td><td>Sequence 10932, A</td></td>	22.9	1031	26 <td>US-10-072-851-10932</td> <td>Sequence 10932, A</td>	US-10-072-851-10932	Sequence 10932, A
21	8 <th>22.9</th> <td>1061</td> <td>28<td>US-10-282-122A-57375</td><td>Sequence 57375, A</td></td>	22.9	1061	28 <td>US-10-282-122A-57375</td> <td>Sequence 57375, A</td>	US-10-282-122A-57375	Sequence 57375, A
22	44 <th>20.0</th> <td>106</td> <td>30<td>US-10-437-963-166960</td><td>Sequence 166960, A</td></td>	20.0	106	30 <td>US-10-437-963-166960</td> <td>Sequence 166960, A</td>	US-10-437-963-166960	Sequence 166960, A
23	7 <th>20.0</th> <td>106</td> <td>20<td>US-09-617-681A-5963</td><td>Sequence 5963, Ap</td></td>	20.0	106	20 <td>US-09-617-681A-5963</td> <td>Sequence 5963, Ap</td>	US-09-617-681A-5963	Sequence 5963, Ap
24	7 <th>20.0</th> <td>106</td> <td>20<td>US-09-688-051-1662</td><td>Sequence 1662, Ap</td></td>	20.0	106	20 <td>US-09-688-051-1662</td> <td>Sequence 1662, Ap</td>	US-09-688-051-1662	Sequence 1662, Ap
25	7 <th>20.0</th> <td>108</td> <td>20<td>US-09-617-681A-5962</td><td>Sequence 5962, Ap</td></td>	20.0	108	20 <td>US-09-617-681A-5962</td> <td>Sequence 5962, Ap</td>	US-09-617-681A-5962	Sequence 5962, Ap
26	7 <th>20.0</th> <td>108</td> <td>19<td>US-09-688-051-1661</td><td>Sequence 1661, Ap</td></td>	20.0	108	19 <td>US-09-688-051-1661</td> <td>Sequence 1661, Ap</td>	US-09-688-051-1661	Sequence 1661, Ap
27	7 <th>20.0</th> <td>109</td> <td>19<td>US-09-513-996A-28293</td><td>Sequence 28293, A</td></td>	20.0	109	19 <td>US-09-513-996A-28293</td> <td>Sequence 28293, A</td>	US-09-513-996A-28293	Sequence 28293, A
28	7 <th>20.0</th> <td>110</td> <td>19<td>US-09-595-328C-256</td><td>Sequence 256, App</td></td>	20.0	110	19 <td>US-09-595-328C-256</td> <td>Sequence 256, App</td>	US-09-595-328C-256	Sequence 256, App
29	7 <th>20.0</th> <td>121</td> <td>20<td>US-09-621-976-4149</td><td>Sequence 4149, Ap</td></td>	20.0	121	20 <td>US-09-621-976-4149</td> <td>Sequence 4149, Ap</td>	US-09-621-976-4149	Sequence 4149, Ap
30	7 <th>20.0</th> <td>121</td> <td>23<td>US-09-834-366-15927</td><td>Sequence 15927, A</td></td>	20.0	121	23 <td>US-09-834-366-15927</td> <td>Sequence 15927, A</td>	US-09-834-366-15927	Sequence 15927, A
31	7 <th>20.0</th> <td>121</td> <td>23<td>US-09-834-366-15927</td><td>Sequence 15927, A</td></td>	20.0	121	23 <td>US-09-834-366-15927</td> <td>Sequence 15927, A</td>	US-09-834-366-15927	Sequence 15927, A
32	7 <th>20.0</th> <td>121</td> <td>31<td>US-09-974-879-143</td><td>Sequence 143, App</td></td>	20.0	121	31 <td>US-09-974-879-143</td> <td>Sequence 143, App</td>	US-09-974-879-143	Sequence 143, App
33	7 <th>20.0</th> <td>121</td> <td>31<td>US-09-147-499-4149</td><td>Sequence 4149, App</td></td>	20.0	121	31 <td>US-09-147-499-4149</td> <td>Sequence 4149, App</td>	US-09-147-499-4149	Sequence 4149, App
34	7 <th>20.0</th> <td>122</td> <td>1<td>PCT-US98-23435-139</td><td>Sequence 15927, A</td></td>	20.0	122	1 <td>PCT-US98-23435-139</td> <td>Sequence 15927, A</td>	PCT-US98-23435-139	Sequence 15927, A
35	7 <th>20.0</th> <td>122</td> <td>17<td>US-09-305-336-142</td><td>Sequence 142, App</td></td>	20.0	122	17 <td>US-09-305-336-142</td> <td>Sequence 142, App</td>	US-09-305-336-142	Sequence 142, App
36	7 <th>20.0</th> <td>122</td> <td>23<td>US-09-818-683-142</td><td>Sequence 142, App</td></td>	20.0	122	23 <td>US-09-818-683-142</td> <td>Sequence 142, App</td>	US-09-818-683-142	Sequence 142, App
37	7 <th>20.0</th> <td>150</td> <td>30<td>US-10-424-599-259484</td><td>Sequence 259484, A</td></td>	20.0	150	30 <td>US-10-424-599-259484</td> <td>Sequence 259484, A</td>	US-10-424-599-259484	Sequence 259484, A
38	7 <th>20.0</th> <td>160</td> <td>19<td>US-09-513-996A-28292</td><td>Sequence 28292, A</td></td>	20.0	160	19 <td>US-09-513-996A-28292</td> <td>Sequence 28292, A</td>	US-09-513-996A-28292	Sequence 28292, A
39	7 <th>20.0</th> <td>161</td> <td>19<td>US-09-595-328C-255</td><td>Sequence 255, App</td></td>	20.0	161	19 <td>US-09-595-328C-255</td> <td>Sequence 255, App</td>	US-09-595-328C-255	Sequence 255, App
40	7 <th>20.0</th> <td>186</td> <td>1<td>PCT-US01-09369-30</td><td>Sequence 30, Appl1</td></td>	20.0	186	1 <td>PCT-US01-09369-30</td> <td>Sequence 30, Appl1</td>	PCT-US01-09369-30	Sequence 30, Appl1
41	7 <th>20.0</th> <td>186</td> <td>1<td>PCT-US98-25512-10</td><td>Sequence 10, Appl1</td></td>	20.0	186	1 <td>PCT-US98-25512-10</td> <td>Sequence 10, Appl1</td>	PCT-US98-25512-10	Sequence 10, Appl1
42	7 <th>20.0</th> <td>186</td> <td>16<td>US-09-203-106-10</td><td>Sequence 10, Appl1</td></td>	20.0	186	16 <td>US-09-203-106-10</td> <td>Sequence 10, Appl1</td>	US-09-203-106-10	Sequence 10, Appl1
43	7 <th>20.0</th> <td>186</td> <td>19<td>US-09-539-330-30</td><td>Sequence 30, Appl1</td></td>	20.0	186	19 <td>US-09-539-330-30</td> <td>Sequence 30, Appl1</td>	US-09-539-330-30	Sequence 30, Appl1
44	7 <th>20.0</th> <td>186</td> <td>21<td>US-09-729-674-30</td><td>Sequence 30, Appl1</td></td>	20.0	186	21 <td>US-09-729-674-30</td> <td>Sequence 30, Appl1</td>	US-09-729-674-30	Sequence 30, Appl1
45	7 <th>20.0</th> <td>193</td> <td>30<td>US-10-437-963-160845</td><td>Sequence 160845, A</td></td>	20.0	193	30 <td>US-10-437-963-160845</td> <td>Sequence 160845, A</td>	US-10-437-963-160845	Sequence 160845, A

ALIGNMENTS

RESULT 1
US-08-842-385-8
Sequence 8, Application US/08842385
GENERAL INFORMATION:
APPLICANT: Russell, John
TITLE OF INVENTION: Colpites, Tracey
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL FOR DETECTING DISEASE OF THE PROSTATE
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/842,385
FILING DATE:
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Porembski, Priscilla E
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 6084.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/937-6365
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 35 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: None
US-08-842-385-8

Query Match 100.0%; Score 35; DB 12; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.8e-27;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SF0SSSTPSTGTGFGSKETPSDDRSQSRHMGES 35
Db 1 SF0SSSTPSTGTGFGSKETPSDDRSQSRHMGES 35

RESULT 2
US-09-991-681-29
Sequence 29, Application US/09991681
GENERAL INFORMATION:
APPLICANT: BILLING-MEDLE, PATRICIA
COHEN, MAURICE
COLPITTS, TRACEY L.
FRIEDMAN, PAULA N.
GORDON, JULIAN
GRANADOS, EDWARD N.
HODGES, STEVEN C.
KLASS, MICHAEL R.
KRATOCHVIL, JON D.
ROBERTS-RAPP, LISA

TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
FOR DETECTING DISEASES OF THE PROSTATE

NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/991,681
FILING DATE: 26-Nov-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/065,383
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6084.US.PI
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 35 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: None
SEQUENCE DESCRIPTION: SEQ ID NO: 29:
US-09-991-681-29

Query Match 100.0%; Score 35; DB 25; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.8e-27;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SF0SSSTPSTGTGFGSKETPSDDRSQSRHMGES 35
Db 1 SF0SSSTPSTGTGFGSKETPSDDRSQSRHMGES 35

RESULT 3
US-08-842-385-6
Sequence 6, Application US/08842385
GENERAL INFORMATION:
APPLICANT: Russell, John
COLPITTS, TRACEY
FRIEDMAN, PAULA N.
GORDON, JULIAN
GRANADOS, EDWARD N.
HODGES, STEVEN C.
KLASS, MICHAEL R.
KRATOCHVIL, JON D.
ROBERTS-RAPP, LISA

TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
FOR DETECTING DISEASE OF THE PROSTATE

NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/842,385
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Porembski, Priscilla E
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 6084.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/937-6365
TELEFAX: 847/938-2623
TELEX:

INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: None
US-08-842-385-6

Query Match 100.0%; Score 35; DB 12; Length 467;
Best Local Similarity 100.0%; Pred. No. 2.3e-26;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SF0SSSTPSTGTGFGSKETPSDDRSQSRHMGES 35
Db 232 SF0SSSTPSTGTGFGSKETPSDDRSQSRHMGES 266

```
RESULT 4
US-09-991-681-27
; Sequence 27, Application US/09991681
; GENERAL INFORMATION:
; APPLICANT: BILLING-MEDL, PATRICIA
; COHEN, MAURICE
; COLPITTS, TRACEY L.
; FRIEDMAN, PAULA N.
; GORDON, JULIAN
; GRANADOS, EDWARD N.
; HODGES, STEVEN C.
; KLAAS, MICHAEL R.
; KRATOCHVIL, JON D.
; ROBERTS-RAPP, LISA
; TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
; FOR DETECTING DISEASES OF THE PROSTATE
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/991,681
; FILING DATE: 26-Nov-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/065,383
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Becker, Cheryl L.
; REGISTRATION NUMBER: 35,441
; REFERENCE/POCKET NUMBER: 6084.US.P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847/935-1729
; TELEFAX: 847/938-2623
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 518 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: None
; SEQUENCE DESCRIPTION: SEQ ID NO: 27:
US-09-991-681-27

Query Match 100.0%; Score 35; DB 25; Length 518;
Best Local Similarity 100.0%; Pred. No. 2.6e-26;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 SFOSSSTPSTGGFGSKETPSDDRSQSRHMGES 35
Db 283 SFOSSSTPSTGGFGSKETPSDDRSQSRHMGES 317

RESULT 5
PCT-US03-01943-44
; Sequence 44, Application PC/TUS0301943
; GENERAL INFORMATION:
; APPLICANT: ORIGENE TECHNOLOGIES INC
; TITLE OF INVENTION: CANCER GENES
; FILE REFERENCE: 3U 9U 901 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/01943
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: US 10/054,935
```

```
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: US 60/356,130
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 10/102,946
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: US 10/117,229
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: US 10/144,198
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: US 10/197,824
; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44
; LENGTH: 1770
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-01943-44

Query Match 100.0%; Score 35; DB 1; Length 1770;
Best Local Similarity 100.0%; Pred. No. 8.6e-26;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SFOSSSTPSTGGFGSKETPSDDRSQSRHMGES 35
Db 1535 SFOSSSTPSTGGFGSKETPSDDRSQSRHMGES 1569

RESULT 6
US-10-144-198-44
; Sequence 44, Application US/10144198
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies Inc
; TITLE OF INVENTION: Regulated Prostate Cance Genes
; FILE REFERENCE: 9U 105 R1
; CURRENT APPLICATION NUMBER: US/10/144,198
; CURRENT FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 44
; LENGTH: 1770
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-198-44

Query Match 100.0%; Score 35; DB 27; Length 1770;
Best Local Similarity 100.0%; Pred. No. 8.6e-26;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SFOSSSTPSTGGFGSKETPSDDRSQSRHMGES 35
Db 1535 SFOSSSTPSTGGFGSKETPSDDRSQSRHMGES 1569

RESULT 7
PCT-US01-08631-40087
; Sequence 40087, Application PC/TUS0108631
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-049
; CURRENT APPLICATION NUMBER: PCT/US01/08631
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 40087
; LENGTH: 1807
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (48)..(62)
OTHER INFORMATION: Fusion glycoprotein F0 domain identified by eMATRIX,
OTHER INFORMATION: accession number PF00523D, p-value=7.188e-10, raw score of 11.39
NAME/KEY: DOMAIN
LOCATION: (941)..(950)
OTHER INFORMATION: Helper component proteinase domain identified by Pfam,
OTHER INFORMATION: accession name Peptidase_C6, E-value=0.0056, Pfam score of 7.9
PCT-US01-08631-40087

Query Match          100.0%; Score 35; DB 1; Length 1807;
Best Local Similarity 100.0%; Pred. No. 8.8e-26;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SF05SSTPSTGFGSGKETPSDDRSQSRHMGES 35
Db 1572 SF05SSTPSTGFGSGKETPSDDRSQSRHMGES 1606

RESULT 8
PCT-US01-42950-495
Sequence 495, Application PC/TUS0142950
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 21272-096
CURRENT APPLICATION NUMBER: PCT/US01/42950
CURRENT FILING DATE: 2001-11-16
PRIOR APPLICATION NUMBER: 09/714,936
PRIOR FILING DATE: 2000-11-17
NUMBER OF SEQ ID NOS: 682
SOFTWARE: PatentIn version 3.0
SEQ ID NO 495
LENGTH: 1839
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US01-42950-495

Query Match          100.0%; Score 35; DB 1; Length 1839;
Best Local Similarity 100.0%; Pred. No. 8.9e-26;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SF05SSTPSTGFGSGKETPSDDRSQSRHMGES 35
Db 1604 SF05SSTPSTGFGSGKETPSDDRSQSRHMGES 1638

RESULT 9
US-10-416-993-495
Sequence 495, Application US/10416993
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 21272-096
CURRENT APPLICATION NUMBER: US/10/416,993
CURRENT FILING DATE: 2003-11-16
PRIOR APPLICATION NUMBER: 09/714,936
PRIOR FILING DATE: 2000-11-17
NUMBER OF SEQ ID NOS: 682
SOFTWARE: PatentIn version 3.0
SEQ ID NO 495
LENGTH: 1839
TYPE: PRT
ORGANISM: Homo sapiens
US-10-416-993-495

Query Match          100.0%; Score 35; DB 30; Length 1839;
Best Local Similarity 100.0%; Pred. No. 8.9e-26;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SF05SSTPSTGFGSGKETPSDDRSQSRHMGES 35
Db 1572 SF05SSTPSTGFGSGKETPSDDRSQSRHMGES 1606

FEATURE:
NAME/KEY: DOMAIN
LOCATION: (48)..(62)
OTHER INFORMATION: Fusion glycoprotein F0 domain identified by eMATRIX,
OTHER INFORMATION: accession number PF00523D, p-value=7.188e-10, raw score of 11.39
NAME/KEY: DOMAIN
LOCATION: (941)..(950)
OTHER INFORMATION: Helper component proteinase domain identified by Pfam,
OTHER INFORMATION: accession name Peptidase_C6, E-value=0.0056, Pfam score of 7.9
PCT-US01-08631-40087

Query Match          100.0%; Score 35; DB 1; Length 1807;
Best Local Similarity 100.0%; Pred. No. 8.8e-26;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SF05SSTPSTGFGSGKETPSDDRSQSRHMGES 35
Db 1572 SF05SSTPSTGFGSGKETPSDDRSQSRHMGES 1606

RESULT 10
PCT-US03-04508-32
Sequence 32, Application PC/TUS0304508
GENERAL INFORMATION:
APPLICANT: IDEC PHARMACEUTICALS
APPLICANT: GATELY, DENNIS
TITLE OF INVENTION: PROSTATE SPECIFIC GENES AND THE USE THEREOF IN DESIGN
FILE REFERENCE: 037003/0301985
CURRENT APPLICATION NUMBER: PCT/US03/04508
CURRENT FILING DATE: 2003-02-19
PRIOR APPLICATION NUMBER: 60/357,140
PRIOR FILING DATE: 2002-02-19
PRIOR APPLICATION NUMBER: 60/396,082
PRIOR FILING DATE: 2002-07-17
PRIOR APPLICATION NUMBER: 60/386,759
PRIOR FILING DATE: 2002-06-10
NUMBER OF SEQ ID NOS: 93
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 32
LENGTH: 1872
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US03-04508-32

Query Match          100.0%; Score 35; DB 1; Length 1872;
Best Local Similarity 100.0%; Pred. No. 9.1e-26;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SF05SSTPSTGFGSGKETPSDDRSQSRHMGES 35
Db 1637 SF05SSTPSTGFGSGKETPSDDRSQSRHMGES 1671

RESULT 11
PCT-US01-08631-40090
Sequence 40090, Application PC/TUS0108631
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 21272-049
CURRENT APPLICATION NUMBER: PCT/US01/08631
CURRENT FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 09/540,217
PRIOR FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: 09/649,167
PRIOR FILING DATE: 2000-08-23
NUMBER OF SEQ ID NOS: 60736
SOFTWARE: Custom
SEQ ID NO 40090
LENGTH: 1982
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (11)..(25)
OTHER INFORMATION: Fusion glycoprotein F0 domain identified by eMATRIX,
OTHER INFORMATION: accession number PF00523D, p-value=7.188e-10, raw score of 11.39
NAME/KEY: DOMAIN
LOCATION: (1065)..(1074)
OTHER INFORMATION: Helper component proteinase domain identified by Pfam,
OTHER INFORMATION: accession name Peptidase_C6, E-value=0.0056, Pfam score of 7.9
PCT-US01-08631-40090

Query Match          100.0%; Score 35; DB 1; Length 1982;
Best Local Similarity 100.0%; Pred. No. 9.6e-26;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SF05SSTPSTGFGSGKETPSDDRSQSRHMGES 35
Db 1572 SF05SSTPSTGFGSGKETPSDDRSQSRHMGES 1606
```



```
Db      1686 SFQSSSTPSTGTFSGKTPSEDPRSQRHMGES 1720
PCT-US03-01943-30
RESULT 12
; Sequence 30, Application PC/TUS0301943
; GENERAL INFORMATION:
; APPLICANT: ORIGENE TECHNOLOGIES INC
; TITLE OF INVENTION: CANCER GENES
; FILE REFERENCE: 3U 9U 901 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/01943
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: US 10/054,935
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: US 60/356,130
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 10/102,946
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: US 10/117,229
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: US 10/144,198
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: US 10/197,824
; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30
; LENGTH: 2221
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-01943-30

Query Match      100.0%; Score 35; DB 1; Length 2221;
Best Local Similarity 100.0%; Pred. No. 1,1e-25;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SFQSSSTPSTGTFSGKTPSEDPRSQRHMGES 35
Db      1986 SFQSSSTPSTGTFSGKTPSEDPRSQRHMGES 2020
PCT-US03-01943-30

RESULT 13
; Sequence 30, Application US/10144198
; GENERAL INFORMATION:
; APPLICANT: OriGene Technologies Inc
; TITLE OF INVENTION: Regulated Prostate Cance Genes
; FILE REFERENCE: 9U 105 R1
; CURRENT APPLICATION NUMBER: US/10/144,198
; CURRENT FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 30
; LENGTH: 2221
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-198-30

Query Match      100.0%; Score 35; DB 27; Length 2221;
Best Local Similarity 100.0%; Pred. No. 1,1e-25;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SFQSSSTPSTGTFSGKTPSEDPRSQRHMGES 35
Db      1986 SFQSSSTPSTGTFSGKTPSEDPRSQRHMGES 2020
PCT-US03-01943-30

RESULT 14
US-10-221-279-7783
; Sequence 7783, Application US/10221279
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
US-10-221-279-7783

; FILE REFERENCE: 21272-046
; CURRENT APPLICATION NUMBER: US/10/221,279
; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 12360
; SOFTWARE: Custom
; SEQ ID NO 7783
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(141)
; OTHER INFORMATION: Xaa = any amino acid or nothing
US-10-221-279-7783

Query Match      57.1%; Score 20; DB 28; Length 141;
Best Local Similarity 100.0%; Pred. No. 9e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      16 GKTPSEDDRSQRHMGES 35
Db      62 GKTPSEDDRSQRHMGES 81
PCT-US03-01943-30

RESULT 15
US-09-134-000-3553
; Sequence 3553, Application US/09134000A
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS
; FILE REFERENCE: GTC-005
; CURRENT APPLICATION NUMBER: US/09/134,000A
; CURRENT FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 6610
; SEQ ID NO 3553
; LENGTH: 270
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000-3553

Query Match      22.9%; Score 8; DB 15; Length 270;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 SESSTPST 11
Db      50 SESSTPST 57
PCT-US03-01943-30

Search completed: July 24, 2003, 12:45:55
Job time : 53.7306 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: July 24, 2003, 12:22:19 ; Search time 1.02489 Seconds
(without alignments)
140.227 Million cell updates/sec

Title: US-09-991-681-29

Perfect score: 35
Sequence: 1 SFQSESTPTGTGFGSKETPSBDDRSQSRHMGES 35

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 41799 seqs, 4106219 residues

Word size: 0

Total number of hits satisfying chosen parameters: 41799

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/2/paa/PCF_NEW_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	ID	Description
1	35	100.0	1872	US-10-367-978-32	Sequence 32, Appl
2	6	17.1	10	US-10-462-850-3684	Sequence 3684, Ap
3	6	17.1	10	US-10-462-850-3685	Sequence 3685, Ap
4	6	17.1	23	US-10-369-060A-111	Sequence 111, Appl
5	6	17.1	183	US-10-243-739A-33	Sequence 33, Appl
6	6	17.1	183	US-10-243-739A-38	Sequence 38, Appl
7	6	17.1	183	US-10-243-739A-39	Sequence 39, Appl
8	6	17.1	183	US-10-243-739A-42	Sequence 42, Appl
9	6	17.1	183	US-10-243-739A-44	Sequence 44, Appl
10	6	17.1	183	US-10-243-739A-45	Sequence 45, Appl
11	6	17.1	183	US-10-243-739A-46	Sequence 46, Appl
12	6	17.1	183	US-10-243-739A-59	Sequence 59, Appl
13	6	17.1	183	US-10-243-739A-60	Sequence 60, Appl
14	6	17.1	183	US-10-243-739A-63	Sequence 63, Appl
15	6	17.1	183	US-10-243-739A-65	Sequence 65, Appl
16	6	17.1	183	US-10-243-739A-71	Sequence 71, Appl
17	6	17.1	183	US-10-465-811-24	Sequence 24, Appl
18	6	17.1	183	US-10-465-811-29	Sequence 29, Appl
19	6	17.1	183	US-10-465-811-33	Sequence 33, Appl
20	6	17.1	183	US-10-465-811-35	Sequence 35, Appl
21	6	17.1	183	US-10-465-811-36	Sequence 36, Appl
22	6	17.1	183	US-10-465-811-37	Sequence 37, Appl
23	6	17.1	183	US-10-465-811-50	Sequence 50, Appl
24	6	17.1	183	US-10-465-811-51	Sequence 51, Appl
25	6	17.1	183	US-10-465-811-54	Sequence 54, Appl
26	6	17.1	183	US-10-465-811-54	Sequence 54, Appl

27	6	17.1	183	US-10-465-811-56	Sequence 56, Appl
28	6	17.1	183	US-10-465-811-62	Sequence 62, Appl
29	6	17.1	183	US-60-470-920-40	Sequence 40, Appl
30	6	17.1	183	US-60-470-920-42	Sequence 42, Appl
31	6	17.1	184	US-10-243-739A-28	Sequence 28, Appl
32	6	17.1	185	US-10-243-739A-77	Sequence 77, Appl
33	6	17.1	185	US-10-465-811-19	Sequence 19, Appl
34	6	17.1	185	US-10-465-811-68	Sequence 68, Appl
35	6	17.1	185	US-10-465-811-97	Sequence 97, Appl
36	6	17.1	197	US-10-243-739A-143	Sequence 143, Appl
37	6	17.1	197	US-10-465-811-71	Sequence 71, Appl
38	6	17.1	212	US-10-243-739A-31	Sequence 31, Appl
39	6	17.1	212	US-10-243-739A-32	Sequence 32, Appl
40	6	17.1	212	US-10-243-739A-34	Sequence 34, Appl
41	6	17.1	212	US-10-243-739A-35	Sequence 35, Appl
42	6	17.1	212	US-10-243-739A-36	Sequence 36, Appl
43	6	17.1	212	US-10-243-739A-40	Sequence 40, Appl
44	6	17.1	212	US-10-243-739A-41	Sequence 41, Appl
45	6	17.1	212	US-10-243-739A-43	Sequence 43, Appl

ALIGNMENTS

```
RESULT 1
US-10-367-978-32
; Sequence 32, Application US/10367978
; GENERAL INFORMATION:
; APPLICANT: GATELY, DENNIS
; TITLE OF INVENTION: PROSTATE SPECIFIC GENES AND THE USE THEREOF IN DESIGN
; FILE REFERENCE: 037003-0301988
; CURRENT APPLICATION NUMBER: US/10/367,978
; PRIOR FILING DATE: 2003-02-19
; PRIOR APPLICATION NUMBER: 60/357,140
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: 60/396,082
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 60/386,759
; PRIOR FILING DATE: 2002-06-10
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 1872
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-367-978-32

Query Match          100.0%; Score 35; DB 6; Length 1872;
Best Local Similarity 100.0%; Pred. No. 4.1e-27;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1 SFQSESTPTGTGFGSKETPSBDDRSQSRHMGES 35
Db      1637 SFQSESTPTGTGFGSKETPSBDDRSQSRHMGES 1671

RESULT 2
US-10-462-850-3684
; Sequence 3684, Application US/10462850
; GENERAL INFORMATION:
; APPLICANT: Proceom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/10/462,850
; PRIOR FILING DATE: 2003-06-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProPatent version 1.0
; SEQ ID NO 3684
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
```

OTHER INFORMATION: sequence located in ATOH1 OR ATH1 at 315-324 and may interact with
OTHER INFORMATION: Sequence 3685 in this patent.
US-10-462-850-3684

Query Match
Best Local Similarity 17.1%; Score 6; DB 6; Length 10;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 TPSTGG 13
DB 2 TPSTGG 7

RESULT 3
US-10-462-850-3685
Sequence 3685, Application US/10462850
GENERAL INFORMATION:
APPLICANT: Proteom Ltd
TITLE OF INVENTION: Complementary peptide ligands from the human genome
FILE REFERENCE: Human patent
CURRENT APPLICATION NUMBER: US/10/462,850
CURRENT FILING DATE: 2003-06-17
NUMBER OF SEQ ID NOS: 4203
SOFTWARE: ProPatent version 1.0
SEQ ID NO 3685
LENGTH: 10
TYPE: PRT
ORGANISM: Homo Sapiens
FEATURE:
OTHER INFORMATION: sequence located in ATOH1 OR ATH1 at 315-324 and may interact with
US-10-462-850-3685

Query Match
Best Local Similarity 17.1%; Score 6; DB 6; Length 10;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 TPSTGG 13
DB 2 TPSTGG 7

RESULT 4
US-10-369-060A-111
Sequence 111, Application US/10369060A
GENERAL INFORMATION:
APPLICANT: Saliberg, Matti
TITLE OF INVENTION: SYNTHETIC PEPTIDES THAT BIND TO THE
FILE REFERENCE: TRIPEP 020CPIC1
CURRENT APPLICATION NUMBER: US/10/369,060A
CURRENT FILING DATE: 2003-02-14
PRIOR APPLICATION NUMBER: 09/839,447
PRIOR FILING DATE: 2001-04-20
PRIOR APPLICATION NUMBER: 09/556,605
PRIOR FILING DATE: 2000-04-21
NUMBER OF SEQ ID NOS: 111
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 111
LENGTH: 23
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Mus musculus
US-10-369-060A-111

Query Match
Best Local Similarity 17.1%; Score 6; DB 6; Length 23;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 RSQSRE 30
DB 15 RSQSRE 20

RESULT 5
US-10-243-739A-33
Sequence 33, Application US/10243739A
GENERAL INFORMATION:
APPLICANT: Bachmann, Martin F.
APPLICANT: Stornil, Tazio
TITLE OF INVENTION: In vivo Activation of Antigen Presenting Cells for Enhancement of
FILE REFERENCE: 1700.0210001
CURRENT APPLICATION NUMBER: US/10/243,739A
CURRENT FILING DATE: 2002-09-16
PRIOR APPLICATION NUMBER: US 60/318,967
PRIOR FILING DATE: 2001-09-14
NUMBER OF SEQ ID NOS: 143
SOFTWARE: PatentIn version 3.1
SEQ ID NO 33
LENGTH: 183
TYPE: PRT
ORGANISM: Hepatitis B virus
US-10-243-739A-33

Query Match
Best Local Similarity 17.1%; Score 6; DB 6; Length 183;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 RSQSRE 30
DB 175 RSQSRE 180

RESULT 6
US-10-243-739A-38
Sequence 38, Application US/10243739A
GENERAL INFORMATION:
APPLICANT: Bachmann, Martin F.
APPLICANT: Stornil, Tazio
TITLE OF INVENTION: In vivo Activation of Antigen Presenting Cells for Enhancement of
FILE REFERENCE: 1700.0210001
CURRENT APPLICATION NUMBER: US/10/243,739A
CURRENT FILING DATE: 2002-09-16
PRIOR APPLICATION NUMBER: US 60/318,967
PRIOR FILING DATE: 2001-09-14
NUMBER OF SEQ ID NOS: 143
SOFTWARE: PatentIn version 3.1
SEQ ID NO 38
LENGTH: 183
TYPE: PRT
ORGANISM: Hepatitis B virus
US-10-243-739A-38

Query Match
Best Local Similarity 17.1%; Score 6; DB 6; Length 183;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 RSQSRE 30
DB 175 RSQSRE 180

RESULT 7
US-10-243-739A-39
Sequence 39, Application US/10243739A
GENERAL INFORMATION:
APPLICANT: Bachmann, Martin F.
APPLICANT: Stornil, Tazio
TITLE OF INVENTION: In vivo Activation of Antigen Presenting Cells for Enhancement of
FILE REFERENCE: 1700.0210001
CURRENT APPLICATION NUMBER: US/10/243,739A
CURRENT FILING DATE: 2002-09-16
PRIOR APPLICATION NUMBER: US 60/318,967
PRIOR FILING DATE: 2001-09-14
NUMBER OF SEQ ID NOS: 143
SOFTWARE: PatentIn version 3.1
SEQ ID NO 39
LENGTH: 183
TYPE: PRT
ORGANISM: Hepatitis B virus
US-10-243-739A-39

Query Match
Best Local Similarity 17.1%; Score 6; DB 6; Length 183;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 RSQSRE 30
DB 175 RSQSRE 180

```

; FILE REFERENCE: 1700.0210001
; CURRENT APPLICATION NUMBER: US/10/243.739A
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/318,967
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 39
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Hepatitis B virus
US-10-243-739A-39

Query Match
Best Local Similarity 17.1%; Score 6; DB 6; Length 183;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 25 RSQSRE 30
Db 175 RSQSRE 180

RESULT 8
US-10-243-739A-42
; Sequence 42, Application US/10243739A
; GENERAL INFORMATION:
; APPLICANT: Bachmann, Martin F.
; APPLICANT: Storni, Tazio
; TITLE OF INVENTION: In vivo Activation of Antigen Presenting Cells for Enhancement of
; FILE REFERENCE: 1700.0210001
; CURRENT APPLICATION NUMBER: US/10/243.739A
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/318,967
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic human Hepatitis B virus core protein gene
US-10-243-739A-42

Query Match
Best Local Similarity 17.1%; Score 6; DB 6; Length 183;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 25 RSQSRE 30
Db 175 RSQSRE 180

RESULT 9
US-10-243-739A-44
; Sequence 44, Application US/10243739A
; GENERAL INFORMATION:
; APPLICANT: Bachmann, Martin F.
; APPLICANT: Storni, Tazio
; APPLICANT: Lechner, Franziska
; TITLE OF INVENTION: In vivo Activation of Antigen Presenting Cells for Enhancement of
; FILE REFERENCE: 1700.0210001
; CURRENT APPLICATION NUMBER: US/10/243.739A
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/318,967
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44
; LENGTH: 183

; TYPE: PRT
; ORGANISM: Hepatitis B virus
US-10-243-739A-45
; Sequence 45, Application US/10243739A
; GENERAL INFORMATION:
; APPLICANT: Bachmann, Martin F.
; APPLICANT: Storni, Tazio
; APPLICANT: Lechner, Franziska
; TITLE OF INVENTION: In vivo Activation of Antigen Presenting Cells for Enhancement of
; FILE REFERENCE: 1700.0210001
; CURRENT APPLICATION NUMBER: US/10/243.739A
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/318,967
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 45
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Hepatitis B virus
US-10-243-739A-45

Query Match
Best Local Similarity 17.1%; Score 6; DB 6; Length 183;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 25 RSQSRE 30
Db 175 RSQSRE 180

RESULT 10
US-10-243-739A-46
; Sequence 46, Application US/10243739A
; GENERAL INFORMATION:
; APPLICANT: Bachmann, Martin F.
; APPLICANT: Storni, Tazio
; APPLICANT: Lechner, Franziska
; TITLE OF INVENTION: In vivo Activation of Antigen Presenting Cells for Enhancement of
; FILE REFERENCE: 1700.0210001
; CURRENT APPLICATION NUMBER: US/10/243.739A
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/318,967
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Hepatitis B virus
US-10-243-739A-46

Query Match
Best Local Similarity 17.1%; Score 6; DB 6; Length 183;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 25 RSQSRE 30
Db 175 RSQSRE 180

RESULT 11
US-10-243-739A-46
; Sequence 46, Application US/10243739A
; GENERAL INFORMATION:
; APPLICANT: Bachmann, Martin F.
; APPLICANT: Storni, Tazio
; APPLICANT: Lechner, Franziska
; TITLE OF INVENTION: In vivo Activation of Antigen Presenting Cells for Enhancement of
; FILE REFERENCE: 1700.0210001
; CURRENT APPLICATION NUMBER: US/10/243.739A
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/318,967
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Hepatitis B virus
US-10-243-739A-46

Query Match
Best Local Similarity 17.1%; Score 6; DB 6; Length 183;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 25 RSQSRE 30
Db 175 RSQSRE 180
```

```

RESULT 12
US-10-243-739A-59
; Sequence 59, Application US/10243739A
; GENERAL INFORMATION:
; APPLICANT: Bachmann, Martin F.
; APPLICANT: Storni, Tazio
; APPLICANT: Lechner, Franziska
; TITLE OF INVENTION: In vivo Activation of Antigen Presenting Cells for Enhancement of
; TITLE OF INVENTION: Immune Responses Induced by Virus Like Particles
; FILE REFERENCE: 1700.0210001
; CURRENT APPLICATION NUMBER: US/10/243,739A
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/318,967
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 59
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Hepatitis B virus
US-10-243-739A-59

Query Match      17.1%; Score 6; DB 6; Length 183;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      25 RSQSRE 30
      |||||
Db      175 RSQSRE 180

RESULT 13
US-10-243-739A-60
; Sequence 60, Application US/10243739A
; GENERAL INFORMATION:
; APPLICANT: Bachmann, Martin F.
; APPLICANT: Storni, Tazio
; APPLICANT: Lechner, Franziska
; TITLE OF INVENTION: In vivo Activation of Antigen Presenting Cells for Enhancement of
; TITLE OF INVENTION: Immune Responses Induced by Virus Like Particles
; FILE REFERENCE: 1700.0210001
; CURRENT APPLICATION NUMBER: US/10/243,739A
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/318,967
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 60
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Hepatitis B virus
US-10-243-739A-60

Query Match      17.1%; Score 6; DB 6; Length 183;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      25 RSQSRE 30
      |||||
Db      175 RSQSRE 180

RESULT 14
US-10-243-739A-63
; Sequence 63, Application US/10243739A
; GENERAL INFORMATION:
; APPLICANT: Bachmann, Martin F.
; APPLICANT: Storni, Tazio
; APPLICANT: Lechner, Franziska
; TITLE OF INVENTION: In vivo Activation of Antigen Presenting Cells for Enhancement of
; TITLE OF INVENTION: Immune Responses Induced by Virus Like Particles

```

```

; FILE REFERENCE: 1700.0210001
; CURRENT APPLICATION NUMBER: US/10/243,739A
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/318,967
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 63
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Hepatitis B virus
US-10-243-739A-63

Query Match      17.1%; Score 6; DB 6; Length 183;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      25 RSQSRE 30
      |||||
Db      175 RSQSRE 180

RESULT 15
US-10-243-739A-65
; Sequence 65, Application US/10243739A
; GENERAL INFORMATION:
; APPLICANT: Bachmann, Martin F.
; APPLICANT: Storni, Tazio
; APPLICANT: Lechner, Franziska
; TITLE OF INVENTION: In vivo Activation of Antigen Presenting Cells for Enhancement of
; TITLE OF INVENTION: Immune Responses Induced by Virus Like Particles
; FILE REFERENCE: 1700.0210001
; CURRENT APPLICATION NUMBER: US/10/243,739A
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/318,967
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 65
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Hepatitis B virus
US-10-243-739A-65

Query Match      17.1%; Score 6; DB 6; Length 183;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      25 RSQSRE 30
      |||||
Db      175 RSQSRE 180

Search completed: July 24, 2003, 12:46:19
Job time : 1.22489 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 12:16:59 ; Search time 4.20205 Seconds
(without alignments)
801.015 Million cell updates/sec

Title: US-09-991-681-29

Perfect score: 35

Sequence: 1 SFSESTPSTGFGFKETPSEDSDRSQSRHMGES 35

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283308 seqs, 96168682 residues

Word size: 0

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: PIR_76.*

1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22.9	8	610	2 T22540	hypothetical prote
2	20.0	7	164	2 T03915	hypothetical prote
3	20.0	7	196	2 T39688	blue copper-bindin
4	20.0	7	196	2 T51838	blue copper-bindin
5	20.0	7	257	2 T48058	RING-H2 zinc finger
6	20.0	7	535	2 T23419	hypothetical prote
7	20.0	7	571	2 AB1274	B. subtilis negati
8	20.0	7	737	1 K1RRC	protein kinase C (
9	20.0	7	737	1 K1RRC	protein kinase C (
10	20.0	7	783	2 T38891	hypothetical prote
11	20.0	7	1693	2 T30867	hypothetical prote
12	20.0	7	2010	2 T28157	erythrocyte membra
13	20.0	7	2212	2 T28157	hypothetical prote
14	20.0	7	2647	2 T28157	hypothetical prote
15	17.1	6	89	2 S78568	snRNP protein SMX4
16	17.1	6	93	2 H75281	hypothetical prote
17	17.1	6	110	2 T42618	hypothetical prote
18	17.1	6	115	2 S09866	hypothetical prote
19	17.1	6	123	2 T07237	hypothetical prote
20	17.1	6	134	2 T07237	hypothetical prote
21	17.1	6	147	2 A42435	large-conductance
22	17.1	6	159	2 S07372	leech antiplatelet
23	17.1	6	162	2 A96537	hypothetical prote
24	17.1	6	168	2 G72584	hypothetical prote
25	17.1	6	172	2 T39026	conserved hypotet
26	17.1	6	181	2 S53178	core antigen - hep
27	17.1	6	183	1 NKVIA2	core antigen - hep
28	17.1	6	183	2 S53214	core antigen - hep
29	17.1	6	183	2 S53221	core antigen - hep
30	17.1	6	183	2 S53232	core antigen - hep

30	6	17.1	183	2 S53247	core antigen - hep
31	6	17.1	183	2 S53260	core antigen - hep
32	6	17.1	183	2 S53267	core antigen - hep
33	6	17.1	183	2 S53270	core antigen - hep
34	6	17.1	183	2 S53286	core antigen - hep
35	6	17.1	183	2 S53129	core antigen - hep
36	6	17.1	183	2 S53143	core antigen - hep
37	6	17.1	183	2 S53149	core antigen - hep
38	6	17.1	183	2 S53152	core antigen - hep
39	6	17.1	183	2 S53166	core antigen - hep
40	6	17.1	183	2 S53169	core antigen - hep
41	6	17.1	183	2 S53175	core antigen - hep
42	6	17.1	183	2 S53186	core antigen - hep
43	6	17.1	183	2 S53194	core antigen - hep
44	6	17.1	183	2 S20755	core antigen - hep
45	6	17.1	183	2 S53181	core antigen - hep

ALIGNMENTS

```

RESULT 1
T22540
hypothetical protein F33B6.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T22540
R:White, S.
Submitted to the EMBL Data Library, October 1996
A:Reference number: Z19578
A:Accession: T22540
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-610 <WLT>
A:Cross-references: EMBL:Z81086; PIDN:CAE03116.1; GSPDB:GN00019; CESP:F53B6.6
A:Experimental source: clone F53B6
C:Genetics:
A:Gene: CESP:F53B6.6
A:Map position: 1
A:Insertions: 38/1; 103/1; 192/2; 239/1; 268/2; 300/3; 532/2; 546/1

Query Match      22.9%; Score 8; DB 2; Length 610;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 SESSTPST 11
DB      481 SESSTPST.488

RESULT 2
T03915
hypothetical protein F32D1.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 08-Oct-1999
C:Accession: T03915
R:Becker, M.; Bradshaw, H.; Kramer, J.
Submitted to the EMBL Data Library, July 1997
A:Description: The sequence of C. elegans cosmid F32D1.
A:Reference number: Z15134
A:Accession: T03915
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-164 <BEC>
A:Cross-references: EMBL:AF016427; NID:G2291228; PIDN:AAE65349.1; PID:G2291230
C:Genetics:
A:Map position: V
A:Insertions: 29/3; 62/3; 98/3
A:Note: F32D1.6

Query Match      20.0%; Score 7; DB 2; Length 164;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 5 ESSTPST 11
|||||
Db 81 ESSTPST 87

RESULT 3

139698
blue copper-binding protein, 20K - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 21-Oct-2002
C:Accession: 139698; S25555
R:Van Gysel, A.; Van Montagu, M.; Inze, D.
Gene 136, 79-85, 1993

A:Title: A negatively light-regulated gene from Arabidopsis thaliana encodes a protein
A:Reference number: 139698; MUID:94124044; PMID:8294044
A:Accession: 139698
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-196 <RES>
A:Cross-references: EMBL:Z15058; NID:916202; PIDN:CAA78771.1; PID:916203
C:Genetics:
A:Gene: bcb
A:Introns: 57/3
C:Superfamily: stercuacyanin
C:Keywords: copper binding; transmembrane protein
F:66,107,112,117/Binding site: copper (His, Cys, His, Gln) #status predicted
F:79-113/Disulfide bonds: #status predicted

Query Match 20.0%; Score 7; DB 2; Length 196;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 STPSTCG 13
|||||
Db 144 STPSTCG 150

RESULT 4

151838
blue copper binding protein homolog [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 21-Oct-2002
C:Accession: T51838
R:Yang, K.Y.; Kim, C.S.; Cho, B.H.

Submitted to the EMBL Data Library, October 1998
A:Description: Characterization of a wound-inducible Arabidopsis gene encoding a protein
A:Reference number: Z25481
A:Accession: T51838
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-196 <YAN>
A:Cross-references: EMBL:Y18227; PIDN:CAA77089.1
C:Genetics:
A:Gene: AMI 32
C:Superfamily: stercuacyanin
C:Keywords: copper binding

Query Match 20.0%; Score 7; DB 2; Length 196;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 STPSTCG 13
|||||
Db 144 STPSTCG 150

RESULT 5

748058
RING-H2 zinc finger protein ATL5 - Arabidopsis thaliana
N:Alternate names: protein F26K9.120
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 08-Dec-2000

C:Accession: T48058
R:Blocker, H.; Mewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.; Quettier, F.; Salanoubat
submitted to the Protein Sequence Database, March 2000
A:Reference number: Z24465
A:Accession: T48058
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-257 <BLO>

A:Cross-references: EMBL:AL162651
A:Experimental source: cultivar Columbia; BAC clone F26K9
C:Genetics:
A:Map position: 3
A:Note: F26K9.120
C:Superfamily: Arabidopsis hypothetical protein F1913.22; RING finger homology
F:109-160/Domain: RING finger homology <RKN>

Query Match 20.0%; Score 7; DB 2; Length 257;
Best Local Similarity 100.0%; Pred. No. 7.6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SESSTPS 10
|||||
Db 194 SESSTPS 200

RESULT 6

723419
hypothetical protein K07F5.14 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T23419
R:Hembry, C.

Submitted to the EMBL Data Library, March 1996
A:Reference number: Z19738
A:Accession: T23419
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-535 <WIL>
A:Cross-references: EMBL:Z70284; PIDN:CAA94285.1; GSPDB:GN00022; CESP:K07F5.14
A:Experimental source: clone K07F5
C:Genetics:
A:Gene: CESP:K07F5.14
A:Map position: 4
A:Introns: 75/3; 112/1; 276/3; 384/3; 422/2; 450/3; 465/1; 508/3

Query Match 20.0%; Score 7; DB 2; Length 535;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 PSTGFS 15
|||||
Db 286 PSTGFS 292

RESULT 7

AB1274
B: subtilis negative regulator of ftsZ ring formation (EzrA) homolog Imo1594 [imported]
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AB1274
R:Glaser, P.; Frangul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fahl, H.; Jones, L.M.; Karet, U.
Science 294, 849-852, 2001
A:Authors: Kretz, J.; Kuhn, M.; Kunz, F.; Kurapkac, G.; Madueno, E.; Maitournam, A.; Matok, C.; Schlueter, T.; Simoes, N.; Tietze, A.; Vazquez-Boland, J.A.; Voss, H.; Weiland, A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AB1274
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-571 <GLA>

A:Cross-references: GB:NC_003210; PIDN:CA99672.1; PID:916411023; GSPDB:GN00177

A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: lmo1594

Query Match 20.0%; Score 7; DB 2; Length 571;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 29 REHNGES 35
Db 449 REHNGES 455

RESULT 8

KIRTC
protein kinase C (EC 2.7.1.-) epsilon - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 11-Jun-1999
C;Accession: B28163; B26408; S00216
R;Ono, Y.; Fujii, T.; Ogita, K.; Kikkawa, U.; Igarashi, K.; Nishizuka, Y.
J. Biol. Chem. 263, 6927-6932, 1988

A;Title: The structure, expression, and properties of additional members of the protein
A;Reference number: A92717; MUID:88198270; PMID:2834397
A;Accession: B28163
A;Molecule type: DNA
A;Residues: 1-737 <ONO>

A;Cross-references: GB:M1831; NID:9206182; PIDN:AAA41872.1; PID:9206183
R;Housey, G.M.; O'Brian, C.A.; Johnson, M.D.; Kirschmeier, P.; Weinstein, I.B.
Proc. Natl. Acad. Sci. U.S.A. 84, 1065-1069, 1987

A;Title: Isolation of cDNA clones encoding protein kinase C: evidence for a protein kinase
A;Reference number: A94145; MUID:87147193; PMID:3469647
A;Accession: B26408

A;Molecule type: mRNA

A;Residues: 397-447 'GORGHNDREDFGSGAET', 467 'LSNPTLLLPDGGPLRQ', 487-545, 'C', 547-636
A;Cross-references: GB:M1553; NID:9206192; PIDN:AAA41877.1; PID:9206193
C;Comment: Protein kinase C epsilon and epsilon appear to be encoded by the same gene a
C;Function:

A;Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonine
A;Note: activity is calcium-independent, phospholipid-dependent, and activated by diacylg
C;Superfamily: protein kinase C delta; protein kinase C zeta-binding repeat homology; pr
C;Keywords: alternative splicing; ATP; autophosphorylation; duplication; phorbol ester B

C
F;156-161/Region: pseudophosphorylation motif
F;170-220/Domain: protein kinase C zinc-binding repeat homology <K21>
F;243-292/Domain: protein kinase C zinc-binding repeat homology <K22>
F;406-668/Domain: protein kinase homology <KIN>

F;414-422/Region: protein kinase ATP-binding motif
F;170,201,204,220/Binding site: zinc (His, Cys, Cys, Cys) #status predicted
F;183,186,209,212/Binding site: zinc (Cys, Cys, His, Cys) #status predicted
F;243,273,276,292/Binding site: zinc (His, Cys, Cys, Cys) #status predicted
F;256,259,281,284/Binding site: zinc (Cys, Cys, His, Cys) #status predicted
F;437,456,532,534/Active site: Lys, Glu, Asp, Lys #status predicted

F;703,710/Binding site: phosphate (Thr) (covalent) (by autophosphorylation) #status pred

Query Match 20.0%; Score 7; DB 1; Length 737;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 PSEDDRS 26
Db 338 PSEDDRS 344

RESULT 9

KIMSC
protein kinase C (EC 2.7.1.-) epsilon - mouse
C;Species: Mus musculus (house mouse)
C;Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 21-Nov-1997
C;Accession: S02270

R;Schaap, D.; Parker, P.J.; Bristol, A.; Kriz, R.; Knopf, J.
FEBS Lett. 243, 351-357, 1989

A;Title: Unique substrate specificity and regulatory properties of PKC-epsilon: a ration
A;Reference number: S02270; MUID:89137541; PMID:2917656

A;Accession: S02270
A;Status: nucleic acid sequence not shown; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-737 <SCH>

A;Function:
A;Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonine
A;Note: activity is calcium-independent, phospholipid-dependent, and activated by diacylg
C;Superfamily: protein kinase C delta; protein kinase C zeta-binding repeat homology; pr
C;Keywords: ATP; autophosphorylation; duplication; phorbol ester binding; phospholipid b

F;156-161/Region: pseudophosphorylation motif
F;170-220/Domain: protein kinase C zinc-binding repeat homology <K21>
F;243-292/Domain: protein kinase C zinc-binding repeat homology <K22>
F;406-668/Domain: protein kinase homology <KIN>

F;414-422/Region: protein kinase ATP-binding motif
F;170,201,204,220/Binding site: zinc (His, Cys, Cys, Cys) #status predicted
F;183,186,209,212/Binding site: zinc (Cys, Cys, His, Cys) #status predicted
F;243,273,276,292/Binding site: zinc (His, Cys, Cys, Cys) #status predicted
F;256,259,281,284/Binding site: zinc (Cys, Cys, His, Cys) #status predicted
F;437,456,532,534/Active site: Lys, Glu, Asp, Lys #status predicted

F;703,710/Binding site: phosphate (Thr) (covalent) (by autophosphorylation) #status predi

Query Match 20.0%; Score 7; DB 1; Length 737;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 PSEDDRS 26
Db 338 PSEDDRS 344

RESULT 10

T38891
hypochemical protein SPAC4H3.11c - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: T38891

R;Murphy, L.; Harris, D.; Bartell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, February 1996
A;Reference number: 221807
A;Accession: T38891

A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-783 <MOR>
A;Cross-references: EMBL:Z69380; PIDN:CAA93350.1; GSPDB:GN00066; SPDB:SPAC4H3.11c
A;Experimental source: strain 972h-; cosmid c4H3

C;Genetics:
A;Gene: SPDB:SPAC4H3.11c
A;Map position: 1

Query Match 20.0%; Score 7; DB 2; Length 783;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 SGETRPS 21
Db 210 SGETRPS 216

RESULT 11

T30867
rho-guanine nucleotide exchange factor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 02-Sep-2000
C;Accession: T30867

R;Gebnik, M.F.B.G.; Poland, M.; Kranenburg, O.; VanHoeck, F.P.G.; Moolenaar, W.H.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z20911
A;Accession: T30867

A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1693 <GB>

A;Cross-references: EMBL:U73199; NID:G1657834; PID:G1657835; PIDN:AA818197.1
A;Experimental source: brain

C:Superfamily: protein kinase C zinc-binding repeat homology
P:652-698/Domain: protein kinase C zinc-binding repeat homology <KZN>

Query Match 20.0%; Score 7; DB 2; Length 1693;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SROSESS 7
DB 1650 SROSESS 1656

RESULT 12

T28157 erythrocyte membrane protein 1 - malaria parasite (Plasmodium falciparum)

C:Species: Plasmodium falciparum
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000

C:Accession: T28157
R:Yang, S.O.; Mollish, W.S.; Gut, J.; Wu, J.; Ahn, J.; Petersen, C.; Fujioka, H.; Aikawa

submitted to the EMBL Data Library, July 1995
A:Description: The molecular cloning and DNA sequence analysis of Plasmodium falciparum

A:Reference number: Z20479
A:Accession: T28157

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA

A:Residues: 1-2212 <YAN>
A:Cross-references: EMBL:U31083; NID:G1517813; PID:G1517814; PIDN:AA06961.1

C:Genetics:
A:Gene: EMP1

Query Match 20.0%; Score 7; DB 2; Length 2212;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 SGETRPS 21
DB 1381 SGETRPS 1387

RESULT 13

T28151 hypothetical protein FC93-varT11-1 - malaria parasite (Plasmodium falciparum) (fragment)

C:Species: Plasmodium falciparum
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000

C:Accession: T28151
R:Hernandez-Rivas, R.; Mattei, D.; Sterkers, Y.; Peterson, D.S.; Wellens, T.E.; Scherf,

Mol. Cell. Biol. 17, 604-611, 1997
A:Title: Expressed var genes are found in Plasmodium falciparum subtelomeric regions.

A:Reference number: Z20483; MUID:97154495; PMID:9001213
A:Accession: T28151

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA

A:Residues: 1-2647 <HER>
A:Cross-references: EMBL:U67959; NID:G1794255; PID:G1809295; PIDN:AA047438.1

A:Experimental source: strain FCQ27/PNG
C:Genetics:
A:Introns: 2158/3

A>Note: FC93-varT11-1

Query Match 20.0%; Score 7; DB 2; Length 2647;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 SGETRPS 21
DB 1382 SGETRPS 1388

RESULT 14

S78568 snRNP protein SMX4 - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YLR438c-a; small nuclear protein SMX4

C:Species: Saccharomyces cerevisiae

C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 19-Apr-2002
C:Accession: S78568

R:Cherry, J.M.
submitted to the EMBL Data Library, August 1997

A:Reference number: S78568
A:Accession: S78568

A:Molecule type: DNA
A:Residues: 1-89 <CHE>

A:Cross-references: EMBL:U21094; GSPDB:GNO0012; MIPS:YLR438c-a.
A:Experimental source: strain S288C (AB972)

C:Genetics:
A:Gene: SGD:LSM3; SMX4; MIPS:YLR438c-a

A:Cross-references: MIPS:YLR438c-a; SGD:S0006434
A:Map position: 12L
C:Keywords: nucleus

Query Match 17.1%; Score 6; DB 2; Length 89;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 TPSEDD 24
DB 78 TPSEDD 83

RESULT 15

H75281 hypothetical protein - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000

C:Accession: H75281
R:White, O.; Eissen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

M. Shen, M.; Vamathevan, J.V.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: H75281

A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-93 <WHI>
A:Cross-references: GB:AE002068; GB:AE000513; NID:G6460186; PIDN:AA011926.1; PID:G6460191

A:Experimental source: strain R1
C:Genetics:
A:Gene: DR2378

A:Map position: 1

Query Match 17.1%; Score 6; DB 2; Length 93;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 SSTRPT 11
DB 36 SSTRPT 41

Search completed: July 24, 2003, 12:27:18
Job time : 5.20205 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 12:15:23 ; Search time 2.306 Seconds
(without alignments)
713.761 Million cell updates/sec

Title: US-09-991-681-29

Sequence: 1 SFQSESTPSTGFGSKETPSEDDRSQSRHMGES 35

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size: 0

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7	20.0	196	1 BCB_ARATH	Q07488 arabidopsis
2	7	20.0	737	1 KPCE_MOUSE	P16054 mus musculus
3	7	20.0	737	1 KPCE_RAT	P09216 ratu musculus
4	7	20.0	783	1 YAVB_SCHPO	Q10218 schizosacch
5	7	20.0	1693	1 RIP2_MOUSE	P97438 mus musculus
6	6	17.1	89	1 LSM3_YEAST	P57743 saccharomyc
7	6	17.1	115	1 UTA1_HCVYA	P16826 human cytom
8	6	17.1	123	1 RRI2_CHLVU	P56354 chlorella v
9	6	17.1	147	1 LAPP_HABOF	Q01747 haementeria
10	6	17.1	159	1 Y131_AGR7	P05680 agrobacteri
11	6	17.1	172	1 YD52_SCHRO	Q10307 schizosacch
12	6	17.1	183	1 CORA_HPBV4	P03150 hepatitis b
13	6	17.1	183	1 CORA_HPBV4	P17391 hepatitis b
14	6	17.1	183	1 CORA_HPBV4	P17392 hepatitis b
15	6	17.1	183	1 CORA_HPBV4	P03146 hepatitis b
16	6	17.1	183	1 CORA_HPBV4	P03147 hepatitis b
17	6	17.1	185	1 CORA_HPBV2	P03148 hepatitis b
18	6	17.1	185	1 CORA_HPBV2	P03149 hepatitis b
19	6	17.1	195	1 CORA_HPBV2	P28178 hepatitis b
20	6	17.1	210	1 HIS7_SYNY3	P48054 synecocyst
21	6	17.1	214	1 CORA_HPBV9	P17099 hepatitis b
22	6	17.1	219	1 BGL5_CAEEL	P17486 caenorhabdi
23	6	17.1	219	1 TRPE_CAUCR	P12289 caulobacter
24	6	17.1	293	1 CHIL1_CHLRE	Q00462 chlamydomon
25	6	17.1	295	1 GRK_DROME	P42287 desophila
26	6	17.1	306	1 UCP1_MESAU	P04575 mesocricetu
27	6	17.1	320	1 CHIX_PEA	P36907 pisum sativi
28	6	17.1	321	1 RT01_YEAST	P10662 saccharomyc
29	6	17.1	329	1 KRAY_LACIA	Q96792 mus musculu
30	6	17.1	358	1 HH2R_MOUSE	P23102 ratu musculu
31	6	17.1	358	1 HH2R_RAT	P23102 ratu musculu
32	6	17.1	363	1 ARI3_NEUCR	Q01356 neurospora
33	6	17.1	371	1 VG07_VACCC	P21028 vaccinia vi

34	6	17.1	371	1 VG07_VARV	P32997 variola vir
35	6	17.1	376	1 AMPC_SERMA	P18539 serixella ma
36	6	17.1	376	1 MID2_YEAST	P36027 saccharomyc
37	6	17.1	387	1 MMK1_MEDSA	Q07176 medicago sa
38	6	17.1	433	1 ELT2_CAEEL	Q10655 caenorhabdi
39	6	17.1	437	1 WTMT_YEAST	Q12363 saccharomyc
40	6	17.1	456	1 DNAL_CHLMU	Q9pk64 chlamydia m
41	6	17.1	456	1 DNAL_CHLTP	Q84252 chlamydia t
42	6	17.1	460	1 DNAL_CHLTP	Q92869 chlamydia p
43	6	17.1	464	1 MADE_PYPAB	Q9v2c0 pyrococcus
44	6	17.1	480	1 HOGT_ARATH	Q9m156 arabidopsis
45	6	17.1	488	1 CYL2_BOVIN	Q28092 bos taurus

ALIGNMENTS

RESULT 1	BCB_ARATH	STANDARD;	PRT;	196 AA.
AC	Q07488: 082664;			
DT	01-FEB-1995 (Rel. 31, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Blue copper protein precursor (blue copper binding protein) (AtBCB).			
GN	BCB OR AM1 32 OR AT5G20230 OR F5024.120.			
OS	Arabidopsis thaliana (Mouse-ear cress).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;			
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsis.			
OX	NCBI_Taxid=3702;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=cv. Columbia, and cv. Columbia K85;			
RX	MEDLINE=94124044; PubMed=8294044;			
RA	van Gysel A., van Montagu M., Inze D.;			
RT	"A negatively light-regulated gene from Arabidopsis thaliana encodes			
RT	a protein showing high similarity to blue copper-binding proteins.";			
RL	Gene 136.79-85(1993).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=cv. Wassilewskija;			
RA	Yang K.Y., Kim C.S., Cho B.H.;			
RT	"Characterization of a wound-inducible Arabidopsis gene encoding a			
RT	protein homologous to blue copper binding proteins.";			
RT	Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=cv. Landsberg erecta;			
RX	MEDLINE=20233824; PubMed=10769227;			
RA	Homma T., Goto K.;			
RT	"The Arabidopsis floral homeotic gene PISTILLATA is regulated by			
RT	discrete cis-elements responsive to induction and maintenance			
RT	signals.";			
RT	Development 127:2021-2030(2000).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=cv. Columbia;			
RX	MEDLINE=21016721; PubMed=11130714;			
RA	Tabata S., Kaneo T., Nakamura Y., Kotani H., Kato T., Asamizu E.,			
RA	Miyajima N., Sasaemoto S., Kimura T., Hosouchi T., Kawashima K.,			
RA	Kohara M., Matsumoto M., Matsuno A., Muraki A., Nakayama S.,			
RA	Nakazaki N., Natsu K., Okumura S., Shindo S., Takeuchi C., Wada T.,			
RA	Matsumoto A., Yamada M., Sato S., de la Bastide M.,			
RA	Huang E., Spiegel U., Gnoj L., O'Shaughnessy A., Preston R.,			
RA	Habermann K., Murray J., Johnson D., Rohlfing T., Nelson J.,			
RA	Stonerberg T., Pepin K., Spieth J., Sekhon M., Armstrong J., Becker M.,			
RA	Belter E., Cordum H., Cordes M., Courtney L., Courtney W., Dante M.,			
RA	Du H., Edwards J., Fryman J., Haekens B., Lamar E., Lattelle P.,			
RA	Leonard S., Meyer R., Mulvaney E., Ozersky P., Riley A., Strommatt C.,			
RA	Wagner-Mehersson C., Wollam A., Yokum W., Bell M., Dedina N.,			
RA	Parnell L., Shah R., Rodriguez M., Hoon See L., Vil D., Baker J.,			
RA	Kirchoff K., Toth K., King L., Bahret A., Miller B., Maira M.,			

RA Martensen R., McCombie W.R., Wilson R.K., Murphy G., Bancroft I.,
 RA Volckaert G., Wandut R., Duesterhoeft A., Stiekema W., Pohl T.,
 RA Entlan K.-D., Terryn N., Hartley N., Bent E., Johnson S.,
 RA Langham S.-A., McCullagh B., Robben J., Grymonprez B., Zimmermann W.,
 RA Rampeger U., Medler H., Balke K., Medler E., Peters S.,
 RA van Staveren M., Dirkse W., Moolman P., Klein lanthorst R.,
 RA Weltzenegger T., Bothe S., Rose W., Hauf J., Benetser S., Hempel S.,
 RA Feldpausch M., Lamberth S., Villarroel R., Giesen U., Ardiles W.,
 RA Bentes O., Lemcke K., Kolesov G., Mayer K.F.X., Rudd S., Schoof H.,
 RA Schueller C., Zaccaria P., Mewes H.-W., Bevan M., Franz P.F.,
 RA "Sequence and analysis of chromosome 5 of the plant Arabidopsis
 RA thaliana".
 RA Nature 408:823-826(2000).
 RL Nature 408:823-826(2000).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
 RT "R1EN Arabidopsis full length cDNA clones (RAFLs) sequenced by the
 RT SSP consortium (salk/Stanford/PGEC)";
 RL Submitted (Aug-2001) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: PROBABLY ACTS AS AN ELECTRON CARRIER.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (probable).
 CC -1- DEVELOPMENTAL STAGE: MAXIMUM LEVELS ARE FOUND IN 35 DAY OLD
 CC PLANTLETS WHEN THE ROSETTE IS MATURE, CONSISTING OF 8-10 FULLY
 CC EXPANDED LEAVES, AND AS THE FLORAL STEM STARTS TO FORM. THIS LEVEL
 CC REMAINS CONSTANT DURING THE FURTHER LIFE SPAN OF THE PLANT.
 CC -1- INDUCTION: BY DARK ADAPTATION. THIS GIVES A 20-FOLD INCREASE IN
 CC EXPRESSION.
 CC -1- SIMILARITY: Contains 1 plastocyanin-like domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, Z15058; CAA78771.1; -;
 DR EMBL, Y18227; CAA77089.1; -;
 DR EMBL, AB035137; BAAG6999.1; -;
 DR EMBL, AF296825; -; NOT ANNOTATED_CDS.
 DR EMBL, AY052681; AAK96585.1; -;
 DR EMBL, AY034986; AAK59491.1; -;
 DR PIR, I39698; I39698.
 DR PIR, T51838; T51838.
 DR HSSP, P29602; 1JBR.
 DR InterPro; IPR000923; BlueCu_1.
 DR InterPro; IPR003245; P1cyanin_1.
 DR Pfam; PF02298; Cu_bind_1.
 DR ProDom; PD003122; p1cyanin_1.
 DR PROSITE; PS00196; COPPER_BLUE; 1.
 KW Electron transport; Copper; Transmembrane; Signal; Metal-binding;
 KW Glycoprotein.
 FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 196 BLUE COPPER PROTEIN.
 FT DOMAIN 23 118 PLASTOCYANIN-LIKE.
 FT TRANSMEM 119 143 POTENTIAL.
 FT TRANSMEM 173 189 POTENTIAL.
 FT DISULFID 79 113 BY SIMILARITY.
 FT METAL 66 66 COPPER (BY SIMILARITY).
 FT METAL 107 107 COPPER (BY SIMILARITY).
 FT METAL 112 112 COPPER (BY SIMILARITY).
 FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 44 44 T -> S (IN REF. 1).
 FT CONFLICT 134 134 P -> L (IN REF. 1).
 FT CONFLICT 142 142 P -> L (IN REF. 1).
 SQ SEQUENCE 196 AA; 20053 MW; 05100B50518F0A56 CRC64;

QY 7 STPGCG 13
 DB 144 STPGTG 150
 ID KCPE_MOUSE STANDARD; PRT; 737 AA.
 AC P16054;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Protein kinase C, epsilon type (BC 2.7.1.1-) (PKC-epsilon).
 GN PKCε OR PKCε OR PKCεA.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89137541; PubMed=2917656;
 RA Schaad D., Parker P.J., Bristol A., Kriz R., Knopf J.;
 RT "Unique substrate specificity and regulatory properties of
 RT PKC-epsilon: a rationale for diversity.";
 RL FEBS Lett. 243:351-357(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=98127436; PubMed=9467942;
 RA Wang Q.J., Ace P., Goodnight J., Blumberg P.M., Mischak H.,
 RA Mushinski J.F.;
 RT "The catalytic domain of PKC-epsilon, in reciprocal PKC-delta and -
 RT epsilon chimeras, is responsible for conferring tumorigenicity to
 RT NIH3T3 cells, whereas both regulatory and catalytic domains of
 RT PKC-epsilon contribute to in vitro transformation.";
 RL Oncogene 16:53-60(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Wheeler D.L.;
 RT Submitted (DEC-2000) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: THIS IS CALCIUM-INDEPENDENT, PHOSPHOLIPID-DEPENDENT,
 CC SRINGE- AND THREONINE-SPECIFIC ENZYME.
 CC -1- FUNCTION: PKC IS ACTIVATED BY DIACYLGLYCEROL WHICH IN TURN
 CC PHOSPHORYLATES A RANGE OF CELLULAR PROTEINS. PKC ALSO SERVES AS
 CC THE RECEPTOR FOR PHORBOL ESTERS, A CLASS OF TUMOR PROMOTERS.
 CC -1- SIMILARITY: Contains 2 zinc-dependent phospho-ester and DAG
 CC binding domains.
 CC -1- SIMILARITY: Contains 1 C2 domain.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC PKC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF028009; AAB84189.1; -;
 DR EMBL; AF325507; AAG53692.1; -;
 DR PIR; S02270; KINSCC.
 DR HSSP; P28867; 1PTQ.
 DR GMD; GO:004699; F:calcium independent protein kinase C activity; IDA.
 DR InterPro; IPR000008; C2.
 DR InterPro; IPR002219; DAG_PE-bind.
 DR InterPro; IPR000961; Pkinase_C.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR Pfam; PF00168; C2; 1.
 DR Pfam; PF00130; DAG_PE-bind; 2.

Query Match 20.0%; Score 7; DB 1; Length 196;
 Best Local Similarity 100.0%; Pred. No. 2.1;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DR Pfam; PF00069; kinase, 1.
 DR Pfam; PF00433; kinase C, 1.
 DR PRINTS; PR00008; DAGEDOMAIN.
 DR PRODOM; PD000001; Prot_kinase; 1.
 DR SMART; SM00109; C1; 2.
 DR SMART; SM00239; C2; 1.
 DR SMART; SM00133; S_TK_X; 1.
 DR SMART; SM00220; S_TK; 1.
 DR PROSITE; PS50004; C2 DOMAIN 2; 1.
 DR PROSITE; PS00479; DAG PE BIND DOM 1; 2.
 DR PROSITE; PS50081; DAG PE BIND DOM 2; 2.
 DR PROSITE; PS50107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR Repeat; ATP-binding; Transferase; Phosphorylation;
 KM Serine/threonine-protein kinase; Phorbol-ester binding; Zinc.
 FT DOMAIN 1 99 C2 DOMAIN.
 FT DOMAIN 170 220 PHORBOL-ESTER AND DAG BINDING 1.
 FT DOMAIN 243 292 PHORBOL-ESTER AND DAG BINDING 2.
 FT NP_BIND 408 668 PROTEIN KINASE.
 FT NP_BIND 414 422 ATP (BY SIMILARITY).
 FT BINDING 437 437 ATP (BY SIMILARITY).
 FT ACT_SITE 532 532 BY SIMILARITY.
 FT MOD_RES 703 703 PHOSPHORYLATION (AUTO-) (POTENTIAL).
 FT MOD_RES 710 710 PHOSPHORYLATION (AUTO-) (POTENTIAL).
 SQ SEQUENCE 737 AA; 83560 MW; 7AEBB8C10C9F57 CRC64;

Query Match 20.0%; Score 7; DB 1; Length 737;
 Best Local Similarity 100.0%; Pred. No. 7.2;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 PSEDDRS 26
 Db 338 PSEDDRS 344

RESULT 3
 ID_KPCE RAT STANDARD; PRT; 737 AA.
 AC P09216;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Protein kinase C, epsilon type (BC 2.7.1.-) (nPKC-epsilon).
 GN PKRCE OR PKCE.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=88198270; PubMed=2834397;
 RA Ono Y., Fujii T., Ogita K., Kikkawa U., Igarashi K., Nishizuka Y.;
 RT "The structure, expression, and properties of additional members of
 the protein kinase C family.";
 RL J. Biol. Chem. 263:6927-6932(1988).
 [2]
 RP SEQUENCE OF 135-297 FROM N.A.
 RX MEDLINE=88083621; PubMed=3691811;
 RA Ono Y., Fujii T., Ogita K., Kikkawa U., Igarashi K., Nishizuka Y.;
 RT "Identification of three additional members of rat protein kinase C
 family: delta-, epsilon- and zeta-subspecies.";
 RL FEBS Lett. 226:125-128(1987).
 CC -1- FUNCTION: THIS IS CALCIUM-INDEPENDENT, PHOSPHOLIPID-DEPENDENT,
 SERINE- AND THREONINE-SPECIFIC ENZYME.
 CC -1- FUNCTION: PKC IS ACTIVATED BY DIACYLGLYCEROL WHICH IN TURN
 PHOSPHORYLATES A RANGE OF CELLULAR PROTEINS. PKC ALSO SERVES AS
 THE RECEPTOR FOR PHORBOL ESTERS, A CLASS OF TUMOR PROMOTERS.
 CC -1- SIMILARITY: Contains 2 zinc-dependent phorbol-ester and DAG
 binding domains.
 CC -1- SIMILARITY: Contains 1 C2 domain.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

CC PKC SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1ab-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL; M18331; AAA1872.1; --
 DR PIR; B28163; KIRTC.
 DR PDB; 1GM1; 25-OCT-01.
 DR InterPro; IPR000008; C2.
 DR InterPro; IPR002219; DAG PE-bind.
 DR InterPro; IPR000961; Kinase C.
 DR InterPro; IPR007719; Prot_kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR Pfam; PF00168; C2; 1.
 DR Pfam; PF00130; DAG PE-bind; 2.
 DR Pfam; PF00069; kinase C, 1.
 DR Pfam; PF00433; kinase C, 1.
 DR PRINTS; PR00008; DAGEDOMAIN.
 DR PRODOM; PD000001; Prot_kinase; 1.
 DR SMART; SM00109; C1; 2.
 DR SMART; SM00239; C2; 1.
 DR SMART; SM00133; S_TK_X; 1.
 DR SMART; SM00220; S_TK; 1.
 DR PROSITE; PS50004; C2 DOMAIN 2; 1.
 DR PROSITE; PS00479; DAG PE BIND DOM 1; 2.
 DR PROSITE; PS50081; DAG PE BIND DOM 2; 2.
 DR PROSITE; PS50107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR Repeat; ATP-binding; Transferase; Phosphorylation;
 KM Serine/threonine-protein kinase; Phorbol-ester binding; Zinc;
 KW 3D-structure.
 FT DOMAIN 1 99 C2 DOMAIN.
 FT DOMAIN 170 220 PHORBOL-ESTER AND DAG BINDING 1.
 FT DOMAIN 243 292 PHORBOL-ESTER AND DAG BINDING 2.
 FT NP_BIND 408 668 PROTEIN KINASE.
 FT NP_BIND 414 422 ATP (BY SIMILARITY).
 FT BINDING 437 437 ATP (BY SIMILARITY).
 FT ACT_SITE 532 532 BY SIMILARITY.
 FT MOD_RES 703 703 PHOSPHORYLATION (AUTO-) (POTENTIAL).
 FT MOD_RES 710 710 PHOSPHORYLATION (AUTO-) (POTENTIAL).
 SQ SEQUENCE 737 AA; 83478 MW; 6AD699FEFDD2659F CRC64;

Query Match 20.0%; Score 7; DB 1; Length 737;
 Best Local Similarity 100.0%; Pred. No. 7.2;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 PSEDDRS 26
 Db 338 PSEDDRS 344

RESULT 4
 ID_YAIB SCHPO STANDARD; PRT; 783 AA.
 AC Q10218;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein C4H3.11c in chromosome 1.
 GN SPAC4H3.11C.
 OS Schizosaccharomyces pombe (Fission Yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OC NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.

```

RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Galliliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgourou J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Church C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Goble L., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornby T., Howarth S., Huckle E.J., Hunt S., Jørgensen K.,
RA Jones K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeil C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton K., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodard J., Wolke G., Aert R., Robben J., Grymoprez B.,
RA Wellens J., Vancraet E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel K., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langner I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Medler H., Wamburt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Leclaire V., Mottier S.,
RA Gallibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Useery D., Barrell B.G., Nurse P.,
RA "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z69380; CA93350.1; -
DR PIR: T38891; T38891.
DR GeneDB: Spombe; SPAC4H3.11c; -.
DR Hypothetical protein.
SQ SEQUENCE 783 AA; 89188 MW; FA2FA539AE5682FE CRC64;

Query Match 20.0%; Score 7; DB 1; Length 783;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 SGERPS 21
DB 210 SGERPS 216

RESULT 5
RIP2_MOUSE STANDARD; PRT; 1693 AA.
ID RIP2_MOUSE
AC P97433;
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Rho-guanine nucleotide exchange factor (Rho-interacting protein 2)
DE (RhoGEF) (RIP2).
GN RGNF OR RHOIP2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN NCB1_TaxID=10090;
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=97344280; PubMed=9199174;
RA Gebbink M.F.B.G., Kranenburg O., Poland M., van Horck F.P.G.,
RA Housa B., Mooleenaar W.H.;
RA "Identification of a novel, putative Rho-specific GDP/GTP exchange

```

```

RT factor and a Rho-binding protein: control of neuronal morphology.";
RL J. Cell Biol. 137:1603-1613(1997).
CC -1- FUNCTION: GUANINE NUCLEOTIDE EXCHANGE FACTOR THAT INTERACTS WITH
CC RHOA, BUT NOT WITH RAC OR CDC42. ACTIVATES RHOA TO PROMOTE
CC CYTOSKELETAL CONTRACTION AND INHIBIT NEURITE OUTGROWTH.
CC -1- TISSUE SPECIFICITY: HIGHLY ENRICHED IN THE BRAIN.
CC -1- SIMILARITY: Contains 1 DBL-homology (DB) domain.
CC -1- SIMILARITY: Contains 1 PH domain.
CC -1- SIMILARITY: Contains 1 zinc-dependent phospho-ester and DAG
CC binding domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U73199; AAB18197.1; -
DR PIR: T30867; T30867.
DR MGD: MGI:1346016; Rgnef.
DR InterPro: IPR002219; DAG_PE-bind.
DR InterPro: IPR001331; GDS_CDC24.
DR InterPro: IPR001849; PH.
DR InterPro: IPR000219; RhoGEF.
DR Pfam: PF00130; DAG_PE-bind; 1.
DR Pfam: PF00169; PH; 1.
DR Pfam: PF00621; RhoGEF; 1.
DR SMART: SM00109; CL; 1.
DR SMART: SM00233; PH; 1.
DR SMART: SM00325; RhoGEF; 1.
DR PROSITE: PS00479; DAG_PE_BIND_DOM_1; 1.
DR PROSITE: PS00081; DAG_PE_BIND_DOM_2; 1.
DR PROSITE: PS00741; DH_1; FALSE_NEG.
DR PROSITE: PS50010; DH_2; 1.
DR PROSITE: PS50003; PH_DOMAIN; 1.
DR Guanine-nucleotide releasing factor; Coiled coil.
KM DOMAIN 99 266 LEO-RICH.
FT DOMAIN 279 310 COILED COIL (POTENTIAL).
FT DOMAIN 315 318 POLY-GLU.
FT DOMAIN 652 698 PHORBOL-ESTER AND DAG BINDING.
FT DOMAIN 846 1041 DH.
FT DOMAIN 1095 1184 PH.
FT DOMAIN 1421 1522 COILED COIL (POTENTIAL).
SQ SEQUENCE 1693 AA; 190325 MW; 4E96087C449FE14C CRC64;

Query Match 20.0%; Score 7; DB 1; Length 1693;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SFQESS 7
DB 1650 SFQESS 1656

RESULT 6
LSM3_YEAST STANDARD; PRT; 89 AA.
ID LSM3_YEAST
AC P57743; Q05176; Q06759;
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE U6 snRNA-associated Sm-like protein Lsm3 (Smx4 protein).
DE LSM3 OR SMX4 OR US53 OR YLR438C-A.
GN Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN NCB1_TaxID=4932;
RP SEQUENCE FROM N.A., IDENTIFICATION, AND FUNCTION.
RX MEDLINE=95262648; PubMed=7744014;
RA Seraphin B.;

```

"Sm and Sm-like proteins belong to a large family: identification of RT proteins of the U6 as well as the U1, U2, U4 and U5 snRNPs.";
 RT EMO J. 14:2089-2099(1995).
 [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN-S286C / AB972;
 RC MEDLINE=97313267; PubMed=9169871;
 RA Johnston M., Hillier L., Riles L., Albermann K., Andre B., Ansgorge W.,
 RA Benes V., Brueckner M., Delius H., Dubois E., Duesterhoeft A.,
 RA Entian K.-D., Floeth M., Goffeau A., Hebling U., Henmann K.,
 RA Heuss-Neitzel D., Hilbert H., Hilger F., Kleine K., Koetter P.,
 RA Louis E.-J., Mesenguy F., Mewes H.-W., Miosga T., Moser D.,
 RA Mueller-Auer S., Nentwich U., Obermayer B., Piravandi E., Pohl T.M.,
 RA Poterelle D., Purnelle B., Rechmann S., Rieger M., Rinke M., Rose M.,
 RA Scharte M., Scherens B., Scholler P., Schwager C., Schwarz S.,
 RA Underwood A.P., Urestarazu L.A., Vandenbol M., Verhasselt P.,
 RA Viereckels F., Voet M., Volckaert G., Voss H., Wandt R., Wedler E.,
 RA Weller H., Zimmermann F.K., Zollner A., Hant U., Hohlseil J.D.;
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII";
 RL Nature 387:87-90(1997).
 [3]
 RN SEQUENCE OF 1-33 FROM N.A.
 RP STRAIN-07173;
 RC MEDLINE=95129901; PubMed=7828914;
 RA Grack H.-R., Grohmann L., Kitakawa M., Goldschmidt-Reisin S.;
 RT "Gene MRP-14, encoding mitochondrial ribosomal protein Yml4, is
 RT indispensable for proper non-respiratory cell functions in yeast.";
 RL Gene 152:107-112(1995).
 [4]
 RN SEQUENCE OF 34-89 FROM N.A.
 RP MEDLINE=88055042; PubMed=2824201;
 RA Degols G.;
 RT "Functional analysis of the regulatory region adjacent to the carb
 RT gene of Saccharomyces cerevisiae. Nucleotide sequence, gene fusion
 RT experiments and cis-dominant regulatory mutation analysis.";
 RL Eur. J. Biochem. 169:193-200(1987).
 [5]
 RN IDENTIFICATION, AND CHARACTERIZATION.
 RP MEDLINE=99359359; PubMed=10428970;
 RA Mayes A.E., Verdore L., Legrain P., Beggs J.D.;
 RT "Characterization of Sm-like proteins in yeast and their association
 RT with U6 snRNA.";
 RL EMO J. 18:4321-4331(1999).
 CC -1- FUNCTION: BINDS SPECIFICALLY TO THE 3'-TERMINAL U-TRACT OF U6
 CC SNRNA.
 CC -1- SUBUNIT: LSM SUBUNITS FORM A HETEROMER WITH A DOUGNOT SHAPE (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: Belongs to the snRNP Sm proteins family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Z30582; CAA83056.1; -;
 CC EMBL; X06790; CAA29948.1; -;
 CC EMBL; U21094; -; NOT ANNOTATED_CDS.
 CC PIR; S78568; S78568.
 CC SGD; S0006434; LSM3.
 CC GO; GO:0017070; P:U6 snRNA binding activity; IPT.
 CC GO; GO:0006371; P:mRNA splicing; IPT.
 CC InterPro; IPR001163; snNP_Sm.
 CC Pfam; PF01423; Sm; 1.
 CC SMART; SM00651; Sm; 1.
 CC Nuclear protein; Ribonucleoprotein; mRNA splicing; mRNA processing;
 CC RNA-binding.
 CC SEQUENCE 89 AA; 10030 MW; 787084F6A5945B7 CRC64;
 Query Match 17.1%; Score 6; DB 1; Length 89;

Best Local Similarity 100.0%; Pred. No. 11;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 19 TPSED 24
 Db 78 TPSEDD 83
 RESULT 7
 ID UTA1_HCMVA STANDARD; PRT; 115 AA.
 AC P16826;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 01-FEB-1991 (Rel. 17, Last annotation update)
 DE Hypothetical protein UTA101.
 GN UTA101.
 OS Human cytomegalovirus (strain AD169).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Cytomegalovirus.
 ON NCBI_TaxID=10360;
 [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=90269039; PubMed=2161319;
 RA Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,
 RA Horsnell T., Hutchinson C.A. III, Kouzarides T., Martignetti J.A.,
 RA Preddie E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;
 RT "Analysis of the protein-coding content of the sequence of human
 RT cytomegalovirus strain AD169.";
 RL Curr. Top. Microbiol. Immunol. 154:125-169(1990).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X17403; CAA35337.1; -;
 CC PIR; S09866; S09866.
 CC Hypothetical protein.
 CC SEQUENCE 115 AA; 12184 MW; 5D3BDC55D4BD904F CRC64;
 Query Match 17.1%; Score 6; DB 1; Length 115;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 21 SEDPRS 26
 Db 55 SEDPRS 60
 RESULT 8
 ID R12_CHLVU STANDARD; PRT; 123 AA.
 AC P56354;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Chloroplast 30S ribosomal protein S12.
 GN RPS12.
 OS Chlorella vulgaris.
 OC Chloroplast; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;
 CC Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;
 CC Chlorellaceae; Chlorella.
 CC NCBI_TaxID=3077;
 [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN-IAM C-27 / Tamiya;
 RC MEDLINE=97303241; PubMed=9159184;
 RA Wakauegi T., Nagai T., Kapoor M., Sugita M., Ito M., Ito S.,
 RA Teudzuki J., Nakashima K., Teudzuki T., Suzuki Y., Hamada A., Ohta T.,

RA Inamura A., Yoshinaga K., Sugiyama M.;
 RT "Complete nucleotide sequence of the chloroplast genome from the
 RT green alga *Chlorella vulgaris*: the existence of genes possibly
 RT involved in chloroplast division."
 RL Proc. Natl. Acad. Sci. U.S.A. 94:5967-5972(1997).
 CC -1- FUNCTION: With S4 and S5 plays an important role in translational
 CC accuracy. Located at the interface of the 30S and 50S subunits (By
 CC similarity).
 CC -1- SUBUNIT: Part of the 30S ribosomal subunit.
 CC -1- SUBCELLULAR LOCATION: Chloroplast.
 CC -1- SIMILARITY: BELONGS TO THE S12P FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, AB001684; BAA57884.1; -.
 CC PIR, T07237; T07237.
 CC HAMAP, MF_00403; -; 1.
 CC InterPro: IPR006032; Ribosomal_S12_23.
 CC InterPro: IPR005679; S12_bact_Chlorpl.
 CC Pfam: PF00164; Ribosomal_S12_1.
 CC PRINTS: PR01034; Ribosomal_S12_1.
 CC PRODOM: PD000576; Ribosomal_S12_23; 1.
 CC TIGRFAMs: TIGR00981; rplB_bact; 1.
 CC PROSITE: PS00055; RIBOSOMAL_S12; 1.
 CC KW Ribosomal protein; RNA-binding; rRNA-binding; Chloroplast.
 CC SEQUENCE 123 AA; 13678 MW; F5DBCA6520664BD CRC64;
 CC
 CC Query Match 17.1%; Score 6; DB 1; Length 123;
 CC Best Local Similarity 100.0%; Pred. No. 15;
 CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 24 DRSQSR 29
 CC Db 109 DRSQSR 114
 CC
 CC RESULT 9
 CC LAMP_HABOF
 CC ID_LAMP_HABOF STANDARD; PRT; 147 AA.
 CC AC Q00147;
 CC DT 01-FEB-1994 (Rel. 28; Created)
 CC DT 01-FEB-1994 (Rel. 28; Last sequence update)
 CC DT 15-SEP-2003 (Rel. 42; Last annotation update)
 CC DE Anti-platelet protein precursor.
 CC LAMP.
 CC OS Haementeria officinalis (Mexican leech).
 CC OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinea; Hirudinea;
 CC OC Rhynchobdellida; Glossiphoniidae; Haementeria.
 CC CX NCBI_TaxID=6410;
 CC RN [1]
 CC RP SEQUENCE FROM N.A., AND SEQUENCE OF 60-91 AND 123-139.
 CC RC TISSUE=Salivary gland;
 CC MEDLINE=9220247; PubMed=1551898;
 CC RA Keller P.M., Schultz L.D., Condra C., Karczewski J., Connolly T.M.;
 RT "An inhibitor of collagen-stimulated platelet activation from the
 RT salivary glands of the *Haementeria officinalis* leech. II. Cloning of
 RT the cDNA and expression."
 RL J. Biol. Chem. 267:6899-6904(1992).
 CC -1- FUNCTION: AN INHIBITOR OF COLLAGEN-STIMULATED PLATELET
 CC AGGREGATION, DENSE GRANULE RELEASE AND SEROTONIN RELEASE.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, M81489; AAA29194.1; -.
 CC DR PDB; 1I8N; 08-AUG-01.
 CC KW Signal; 3D-structure.
 CC FT SIGNAL 1 21
 CC CHAIN 22 147 POTENTIAL.
 CC SEQUENCE 147 AA; 15908 MW; 75A5511374A4E42E CRC64;
 CC
 CC Query Match 17.1%; Score 6; DB 1; Length 147;
 CC Best Local Similarity 100.0%; Pred. No. 18;
 CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 8 TPSTGG 13
 CC Db 45 TPSTGG 50
 CC
 CC RESULT 10
 CC Y131_AGR7
 CC ID_Y131_AGR7 STANDARD; PRT; 159 AA.
 CC AC P05680;
 CC DT 01-NOV-1988 (Rel. 09; Created)
 CC DT 01-NOV-1988 (Rel. 09; Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41; Last annotation update)
 CC DE Insertion element IS136 hypothetical 16.9 kDa protein.
 CC OS Agrobacterium tumefaciens (strain 137).
 CC OG Plasmid pTi137.
 CC OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 CC OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
 CC CX NCBI_TaxID=176300;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=A208;
 CC RX MEDLINE=86312927; PubMed=3016677;
 CC RA Vanderleyden J., Desair J., de Weirman C., Michiels K.,
 RA van Gool A.P., Chilton M.-D., Jen G.C.;
 RT "Nucleotide sequence of an insertion sequence (IS) element identified
 RT in the T-DNA region of a spontaneous variant of the Ti-plasmid
 RT pTi137."
 CC Nucleic Acids Res. 14:6699-6709(1986).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, X04282; CAA27829.1; -.
 CC DR PIR; S07372; S07372.
 CC KW Hypothetical protein; Plasmid; Transposable element.
 CC SEQUENCE 159 AA; 16934 MW; EF0158EC821EE6B CRC64;
 CC
 CC Query Match 17.1%; Score 6; DB 1; Length 159;
 CC Best Local Similarity 100.0%; Pred. No. 19;
 CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 6 SSTPST 11
 CC Db 133 SSTPST 138
 CC
 CC RESULT 11
 CC YD52_SCHPO
 CC ID_YD52_SCHPO STANDARD; PRT; 172 AA.
 CC AC Q10307;
 CC DT 01-OCT-1996 (Rel. 34; Created)
 CC DT 01-OCT-1996 (Rel. 34; Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41; Last annotation update)
 CC DE Hypothetical protein C6C3.02c in chromosome I.


```

GN SPAC6C3.02C.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brookes K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feldwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jags K.,
RA James K., Jones L., Jones M., Leachter S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Voiclaert G., Aert R., Robben J., Grymptier B.,
RA Weijens I., Vanstreels E., Rieger W., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Medler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mortier S.,
RA Lucas M., Roche M., Galliard C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Love T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Uesery D., Barrell B.G., Nurse P.;
RA "The genome sequence of Schizosaccharomyces pombe.";
RT Nature 415:871-880(2002).
RL
CC -!- SIMILARITY: SOME, TO YEAST YMR002W.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sb-sib.ch).
CC -----
DR EMBL; Z69731, CA93615.1; -.
DR PIR; T39026; T39026.
DR GeneDB; SPombe; SPAC6C3.02c; -.
DR InterPro; IPR007017; DUF657.
DR Pfam; PF04933; DUF657; 1.
KW Hypothetical protein.
SQ
SEQUENCE 172 AA; 17442 MW; FA0402DDJFC69FA9A CRC64;

```

```

Query Match 17.1%; Score 6; DB 1; Length 172;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 11 TGGFSG 16
DB 83 TGGFSG 88

```

```

RESULT 12
CORR_HPBV4 STANDARD; PRT; 183 AA.
AC P03150; P03151;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE Core antigen.
OS Hepatitis B virus (subtype adr4),
GN C.
OS Hepatitis B virus (subtype adr4),

```

```

OS Hepatitis B virus (subtype adr), and
OS Hepatitis B virus (subtype adw / strain Indonesia/pIDW420).
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10409, 106820, 10412;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=ADR;
RX MEDLINE=83168919; PubMed=6300776;
RA Ono Y., Ono H., Sasada R., Igarashi K., Sugino Y., Nishio K.;
RT "The complete nucleotide sequences of the cloned hepatitis B virus
RT DNA; subtype adr and adw.";
RL Nucleic Acids Res. 11:1747-1757(1983).
RN
RP SEQUENCE FROM N.A.
RC STRAIN=ADR4;
RX MEDLINE=83246570; PubMed=6306594;
RA Fujiyama A., Miyahara A., Nozaki C., Yoneyama T., Ohmoto N.,
RA Matsubara K.;
RT "Cloning and structural analyses of hepatitis B virus DNAs, subtype
RT adr.";
RL Nucleic Acids Res. 11:4601-4610(1983).
RN
RP SEQUENCE FROM N.A.
RC STRAIN=ADM;
RX MEDLINE=89010694; PubMed=3171552;
RA Okamoto H., Tsuda F., Sakugawa H., Sastroseowignjo R.I., Imai M.,
RA Miyakawa Y., Mayumi M.;
RT "Typing hepatitis B virus by homology in nucleotide sequence:
RT comparison of surface antigen subtypes.";
RL J. Gen. Virol. 69:2575-2583(1988).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sb-sib.ch).
CC -----
DR EMBL; V00867; NOT_ANNOTATED_CDS.
DR EMBL; X01587; CA25745.1; -.
DR EMBL; D00331; NOT_ANNOTATED_CDS.
DR InterPro; IPR002006; Hepatitis core.
DR Pfam; PF00906; Hepatitis_core; 1.
KW Core protein; Repeat.
FT REPEAT 162 169
FT REPEAT 170 177
SQ
SEQUENCE 183 AA; 21095 MW; ED2DA1DB07FB596D CRC64;

```

```

Query Match 17.1%; Score 6; DB 1; Length 183;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 25 RSQSRE 30
DB 175 RSQSRE 180

```

```

RESULT 13
CORR_HPBV4 STANDARD; PRT; 183 AA.
AC P17391;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE Core antigen.
OS Hepatitis B virus (subtype adw / strain Japan/pJDM233).
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10413;
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=89010694; PubMed=3171552;
RA Okamoto H., Tsuda F., Sakugawa H., Sastroseowignjo R.I., Imai M.,

```

```

RA Miyakawa Y., Mayumi M.;
RT "Typing hepatitis B virus by homology in nucleotide sequence:
RL comparison of surface antigen subtypes.",
J. Gen. Virol. 69:2575-2583(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D00329; -; NOT ANNOTATED CDS.
DR InterPro; IPR002006; Hepatitis_core.
DR Pfam; PF00906; Hepatitis_core; 1.
KW Core protein; Repeat.
FT REPEAT 162 169
FT REPEAT 170 177
SQ SEQUENCE 183 AA; 21224 MW; 9FDD6B5F5AF5E160 CRC64;

Query Match
Best Local Similarity 17.1%; Score 6; DB 1; Length 183;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 RSQSR 30
Db 175 RSQSR 180

RESULT 14
COR_A HPBV STANDARD; PRT; 183 AA.
ID COR_A HPBV
AC P17392;
DT 01-AUG-1990 (Rel. 15; Created)
DT 01-AUG-1990 (Rel. 15; Last sequence update)
DT 01-FEB-1991 (Rel. 17; Last annotation update)
DE Core antigen.
OS Hepatitis B virus (subtype adw / strain Okinawa/PDDW282).
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10415;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89010694; PubMed=3171552;
RA Okamoto H., Tsuda F., Sakugawa H., Sastrosowijanto R.I., Imai M.,
RA Miyakawa Y., Mayumi M.;
RT "Typing hepatitis B virus by homology in nucleotide sequence:
RT comparison of surface antigen subtypes.",
J. Gen. Virol. 69:2575-2583(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D00330; -; NOT ANNOTATED CDS.
DR InterPro; IPR002006; Hepatitis_core.
DR Pfam; PF00906; Hepatitis_core; 1.
KW Core protein; Repeat.
FT REPEAT 162 169
FT REPEAT 170 177
SQ SEQUENCE 183 AA; 21081 MW; 66BDB2633122930C CRC64;

Query Match
Best Local Similarity 17.1%; Score 6; DB 1; Length 183;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 RSQSR 30
Db 175 RSQSR 180

```

```

RESULT 15
COR_A HPBV STANDARD; PRT; 183 AA.
ID COR_A HPBV
AC P03146;
DT 21-JUL-1986 (Rel. 01; Created)
DT 21-JUL-1986 (Rel. 01; Last sequence update)
DT 28-FEB-2003 (Rel. 41; Last annotation update)
DE Core antigen.
GN C.
OS Hepatitis B virus (subtype ayw).
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10418;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81012091; PubMed=399327;
RA Galibert F., Mandart E., Flouquet F., Tiollais P., Charnay P.;
RT "Nucleotide sequence of the hepatitis B virus genome (subtype ayw)
RT cloned in E. coli.",
RL Nature 281:646-650(1979).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Latvia;
RX MEDLINE=85204397; PubMed=3996597;
RA Bichko V., Pushko P., Dreilina D., Pumpen P., GREN E.;
RT "Subtype ayw variant of hepatitis B virus. DNA primary structure
RT analysis.",
RL FEBS Lett. 185:208-212(1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V01460; CA24706.1; ALT_INIT.
DR EMBL; X02496; -; NOT ANNOTATED CDS.
DR InterPro; IPR002006; Hepatitis_core.
DR Pfam; PF00906; Hepatitis_core; 1.
KW Core protein; Repeat.
FT REPEAT 162 169
FT REPEAT 170 177
FT VARIANT 33 33 T -> N (IN STRAIN LATVIA).
FT VARIANT 80 80 A -> I (IN STRAIN LATVIA).
SQ SEQUENCE 183 AA; 21116 MW; E0D9D9763F24E958 CRC64;

Query Match
Best Local Similarity 17.1%; Score 6; DB 1; Length 183;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 RSQSR 30
Db 175 RSQSR 180

```

Search completed: July 24, 2003, 12:22:09
Job time : 5.306 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 12:16:28 ; Search time 10.7613 Seconds
(without alignments)
839.285 Million cell updates/sec

Title: us-09-991-681-29

Perfect score: 35
Sequence: 1 SFOSSSTPSTGFGSKETPSDDRSQSRHMGES 35

Scoring table:
Gap: 60.0, Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size: 0

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_protein:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	100.0	589	4 Q8N4Y4	Q8N4Y4 homo sapien
2	35	100.0	592	4 Q96CH9	Q96CH9 homo sapien
3	35	100.0	1770	4 Q9ULH6	Q9ULH6 homo sapien
4	8	22.9	610	5 Q9XV2	Q9XV2 caenorhabdi
5	7	20.0	88	17 Q9HNM3	Q9HNM3 halobacteri
6	7	20.0	164	5 Q16292	Q16292 caenorhabdi
7	7	20.0	205	17 Q9HTH4	Q9HTH4 thermoplasma
8	7	20.0	208	10 Q8H6Z0	Q8H6Z0 phytophthor
9	7	20.0	234	16 Q8EWP2	Q8EWP2 mycoplasma
10	7	20.0	257	10 Q9XFE4	Q9XFE4 arabidopsis
11	7	20.0	257	10 Q9LZJ6	Q9LZJ6 arabidopsis
12	7	20.0	309	17 Q8TLM7	Q8TLM7 methanocarc
13	7	20.0	341	5 Q8MR74	Q8MR74 drosophila
14	7	20.0	342	16 Q8NSR7	Q8NSR7 corynebacte
15	7	20.0	422	5 Q21298	Q21298 caenorhabdi
16	7	20.0	422	12 Q9IMS8	Q9IMS8 cherry molt

17	7	20.0	457	3 Q8TFY2	Q8TFY2 aspergillus
18	7	20.0	571	16 Q8Y6T8	Q8Y6T8 listeria mo
19	7	20.0	613	5 Q8SXN5	Q8SXN5 drosophila
20	7	20.0	761	2 Q9F7S9	Q9F7S9 uncultured
21	7	20.0	976	5 Q9VX63	Q9VX63 drosophila
22	7	20.0	1310	5 Q9VVF8	Q9VVF8 drosophila
23	7	20.0	1625	12 Q55597	Q55597 garlic viru
24	7	20.0	2212	5 Q94657	Q94657 plasmodium
25	7	20.0	2647	5 P90580	P90580 plasmodium
26	7	20.0	2781	4 Q9UGS2	Q9UGS2 homo sapien
27	6	17.1	49	5 Q9N367	Q9N367 caenorhabdi
28	6	17.1	58	11 Q8R8X8	Q8R8X8 mus musculu
29	6	17.1	93	16 Q9RRV7	Q9RRV7 dinococcus
30	6	17.1	96	12 Q9Y042	Q9Y042 hepatitis b
31	6	17.1	96	12 Q9Y0M3	Q9Y0M3 hepatitis b
32	6	17.1	105	10 Q40111	Q40111 lycopersico
33	6	17.1	110	12 Q39311	Q39311 equine hepr
34	6	17.1	116	17 Q8TK93	Q8TK93 methanocarc
35	6	17.1	118	16 Q8FP73	Q8FP73 leptospira
36	6	17.1	120	5 Q8IFH1	Q8IFH1 trypanosoma
37	6	17.1	123	12 Q55526	Q55526 equine hepr
38	6	17.1	125	11 Q8CPT8	Q8CPT8 mus musculu
39	6	17.1	129	11 Q8BWA6	Q8BWA6 mus musculu
40	6	17.1	130	12 Q8QP24	Q8QP24 hepatitis b
41	6	17.1	130	12 Q8QP11	Q8QP11 hepatitis b
42	6	17.1	130	12 Q8QP22	Q8QP22 hepatitis b
43	6	17.1	130	12 Q8QP19	Q8QP19 hepatitis b
44	6	17.1	130	12 Q8QP32	Q8QP32 hepatitis b
45	6	17.1	130	12 Q8QP23	Q8QP23 hepatitis b

ALIGNMENTS

RESULT 1

Q8N4Y4 ID Q8N4Y4 PRELIMINARY; PRT; 589 AA.
AC Q8N4Y4; 01-OCT-2002 (TREMURel. 22, Created)
DT 01-OCT-2002 (TREMURel. 22, Last sequence update)
DE 01-OCT-2002 (TREMURel. 22, Last annotation update)
DE Similar to KIAA1244 protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strauberg R.;
DR EMBL, BC033191; AAH33191.1; -.
FT NON_TER
SQ SEQUENCE 589 AA; 66086 MW; 8041BEA348DE65F7 CRC64;

Query March 100.0%; Score 35; DB 4; Length 589;
Best Local Similarity 100.0%; Pred. No. 5.3e-29;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SFOSSSTPSTGFGSKETPSDDRSQSRHMGES 35
DB 354 SFOSSSTPSTGFGSKETPSDDRSQSRHMGES 388

RESULT 2

Q96CH9 ID Q96CH9 PRELIMINARY; PRT; 592 AA.
AC Q96CH9; 01-DEC-2001 (TREMURel. 19, Created)
DT 01-DEC-2001 (TREMURel. 19, Last sequence update)
DT 01-OCT-2002 (TREMURel. 22, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC014227; AAH14227.1; -.
KW Hypothetical protein.
FT NON TER
SQ SEQUENCE 592 AA; 66400 MW; FA41E807B0DF47B5 CRC64;

Query Match
Best Local Similarity 100.0%; Score 35; DB 4; Length 592;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFOSSTPTGTGFGSKETPSDDRSQSRHHMGES 35
DB 357 SFOSSTPTGTGFGSKETPSDDRSQSRHHMGES 391

RESULT 3
O9ULH6 PRELIMINARY; PRT; 1770 AA.
AC O9ULH6; O96P46;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE BIG3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Hong W.;
RT "K1A1244 as a novel distantly related member (BIG3) of the BIG1/Sec7P
RT subfamily of ARF GEFs."
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF413080; AAL04174.1; -.
DR InterPro; IPR000904; Sec7.
DR SMART; SM00222; Sec7; 1.
KW Hypothetical protein.
SQ SEQUENCE 1770 AA; 195845 MW; SE996E36A692B4 CRC64;

Query Match
Best Local Similarity 100.0%; Score 35; DB 4; Length 1770;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFOSSTPTGTGFGSKETPSDDRSQSRHHMGES 35
DB 1535 SFOSSTPTGTGFGSKETPSDDRSQSRHHMGES 1569

RESULT 4
O9XVNZ PRELIMINARY; PRT; 610 AA.
AC O9XVNZ;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE F53B6.6 protein.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA White S.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RN [2]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=99063613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL; Z81086; CAB03116.1; -.
DR WormPeP; F53B6.6; CE10900.
DR InterPro; IPR001507; Endoglin/CD105.
DR Pfam; PF00100; zona_pellucida; 1.
DR SMART; SM00241; ZP; 1.
SQ SEQUENCE 610 AA; 68585 MW; 09819DC5C3741329 CRC64;

Query Match
Best Local Similarity 100.0%; Score 8; DB 5; Length 610;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 SESSPTPT 11
DB 481 SESSPTPT 488

RESULT 5
O9HHM3 PRELIMINARY; PRT; 88 AA.
AC O9HHM3;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE Vng6325h.
GN VNG6325H.
OS Halobacterium sp. (strain NRC-1).
OC Archaea; Euryarchaeota; Halobacteriia; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahaitas G.G., Bergquist B., Pan M.,
RA Shukla H.D., Laaky S.R., Baliga N.S., Thorsson V., Sprogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Weir R., Goo Y.A.,
RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonki P.E., Krebs M.P., Angvine C.M., Dale H.,
RA Isenberger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT "Genome sequence of Halobacterium species NRC-1."
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR EMBL; AE005162; AAC20955.1; -.
KW Plasmid; Complete proteome.
SQ SEQUENCE 88 AA; 9956 MW; 01B048AEB30B6E9 CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 17; Length 88;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 23 DDRSQSR 29
DB 13 DDRSQSR 19

RESULT 6
O16292 PRELIMINARY; PRT; 164 AA.
AC O16292;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE F32D1.6 protein.
GN F32D1.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

```

OC Rhaditidae; Peloderinae; Caenorrhaditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode *C. elegans*: a platform for investigating biology. The *C. elegans* Sequencing Consortium."; Science 282:2012-2018(1998).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=Bristol N2;
 RC Becker M., Bradshaw H., Kramer J.;
 RA "The sequence of *C. elegans* cosmid F32D1."; Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF016427; AA865349.1; -
 DR WormPep; F32D1.6; CE09870.
 SO SEQUENCE 164 AA; 18160 MW; 33805DF834643AF0 CRC64;
 Query Match 20.0%; Score 7; DB 5; Length 164;
 Best Local Similarity 100.0%; Pred. No. 9.9;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 5 ESTPST 11
 Db 81 ESTPST 87

RESULT 7
 ID Q9H1H4 PRELIMINARY; PRT; 205 AA.
 AC Q9H1H4;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Hypothetical protein T91365.
 GN T91365.
 OS Thermoplasma acidophilum.
 OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
 CC Thermoplasmataceae; Thermoplasma.
 OX NCBI_TaxID=2303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DSM 1728;
 RX MEDLINE=20479972; PubMed=11029001;
 RA Ruopp A., Groll W., Santos-Martinez M.-L., Koretke K.K., Volker C., Mees H.-W., Frishman D., Stocker S., Juppas A.N., Baumeister W.;
 RT "The genome sequence of the thermophilic scavenger *Thermoplasma acidophilum*"; Nature 407:508-513(2000).
 RL Nucleur 407:508-513(2000).
 DR EMBL; AL445067; CAC12486.1; -
 DR HSSP; P28692; 1E10.
 DR InterPro; IPR002877; FEU.
 DR InterPro; IPR000051; SAM_bind.
 DR Pfam; PF01728; FtsJ; 1.
 KW Hypothetical protein; Complete proteome.
 SO SEQUENCE 205 AA; 23194 MW; 96EB170B8CD6D8B1 CRC64;
 Query Match 20.0%; Score 7; DB 17; Length 205;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 13 GFSGET 19
 Db 193 GFSGET 199

RESULT 8
 ID Q8H6Z0 PRELIMINARY; PRT; 208 AA.
 AC Q8H6Z0;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Hypothetical protein.
 OS Phytophthora infestans (Potato late blight fungus).
 CC Eukaryota; stramenopiles; Oomycetes; Pythales; Pythaceae;
 OC Phytophthora
 OX NCBI_TaxID=4787;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DDR7602;
 RA Torto T.A., Styer A., Kamoun S.;
 RT "EST Mining and Functional Expression Assays Identify Extracellular Elitor Proteins from *Phytophthora*."; Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF424681; AAN31506.1; -
 KW Hypothetical protein.
 SO SEQUENCE 208 AA; 23390 MW; B5292FC6BEC661 CRC64;
 Query Match 20.0%; Score 7; DB 10; Length 208;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 18 ETPSEDD 24
 Db 49 ETPSEDD 55

RESULT 9
 ID Q8EWP2 PRELIMINARY; PRT; 234 AA.
 AC Q8EWP2;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Conserved hypothetical protein.
 GN MYPE10.
 OS Mycoplasma penetrans.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=28227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HF-2;
 RX MEDLINE=22354719; PubMed=1246555;
 RA Sasaki Y., Ishikawa J., Yamashita A., Oshima K., Kenri T., Furuya K., Yoshino C., Horino A., Shiba T., Sasaki T., Hattori M.;
 RT "The complete genomic sequence of *Mycoplasma penetrans*, an intracellular bacterial pathogen in humans."; Nucleic Acids Res. 30:5293-5300(2002).
 RL Nucleic Acids Res. 30:5293-5300(2002).
 DR EMBL; AP004170; BAQ43952.1; -
 KW Hypothetical protein; Complete proteome.
 SO SEQUENCE 234 AA; 27537 MW; B97CE9F71B18AE3 CRC64;
 Query Match 20.0%; Score 7; DB 16; Length 234;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 12 GFSGKE 18
 Db 189 GFSGKE 195

RESULT 10
 ID Q9XF64 PRELIMINARY; PRT; 257 AA.
 AC Q9XF64;
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)

```

DE RING-H2 zinc finger protein ATL5.
GN ATL5.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxId=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=9408259; PubMed=10480382;
RX Salinas-Mondragon R.B., Garciduenas-Pina C., Guzman P.;
RT "Early elicitor induction in members of a novel multigene family
RT coding for highly related RING-H2 proteins in Arabidopsis thaliana.";
RL Plant Mol. Biol. 40:579-590(1999).
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL; AF132015; RAD3583.1; -.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF000097; zfc3HC4_1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PSS0089; ZF_RING_2; 1.
KW Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 257 AA; 28608 MW; 078CEP8CEC928C96 CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 10; Length 257;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SESSSTPS 10
DB 194 SESSSTPS 200

RESULT 11
Q9LZJ6 PRELIMINARY; PRT; 257 AA.
AC Q9LZJ6;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE RING-H2 zinc finger protein ATL5.
GN F26K9.120.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxId=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bloecher H., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X., Quettier F.,
RA Salenoudat M.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL; AL162651; CAB83119.1; -.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF000097; zfc3HC4_1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PSS0089; ZF_RING_2; 1.
KW Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 257 AA; 28592 MW; B6B7595DFP528431 CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 10; Length 257;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SESSSTPS 10
DB 194 SESSSTPS 200

RESULT 12
Q8TLM7 PRELIMINARY; PRT; 309 AA.
AC Q8TLM7;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Carbon-nitrogen hydrolase.
GN MA3009.
OS Methanoscaccia acetivorans.
OC Archaea; Euryarchaeota; Methanococci; Methanoscarchinales;
OC Methanoscarchiaceae; Methanoscarchina.
OX NCBI_TaxId=2214;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=1193238;
RA Galagan J.E., Nusbaum C., Roy A., Erdilazi M.G., Macdonald P.,
RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
RA Allen N., Naylor J., Stange-Thomann N., DeArrellano K., Johnson R.,
RA Linton L., McEwan P., McKernan K., Talamas J., Titrrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Graham D.A., Guss A.M.,
RA Hedderich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;
RT "The genome of Methanoscaccia acetivorans reveals extensive metabolic
RT and physiological diversity.";
RL Genome Res. 12:532-542(2002).
DR EMBL; AE011003; AAM06382.1; -.
DR InterPro; IPR003010; NClase/CNhydase.
DR Pfam; PF00795; CN_hydrolase; 1.
KW Hydrolase; Complete proteome.
SQ SEQUENCE 309 AA; 34100 MW; EGABA87F3FFEA41 CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 17; Length 309;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 GFSGKET 19
DB 117 GFSGKET 123

RESULT 13
Q8MR74 PRELIMINARY; PRT; 341 AA.
AC Q8MR74;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE AT3411p (Fragment).
GN CG7692.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Stapleton M., Brokslein P., Hong L., Agbayan A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Fiske E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragosa V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Cealiker S.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY122081; AAM52593.1; -.
DR FlyBase; FBgn0036714; CG7692.
FT NON TER
SQ SEQUENCE 341 AA; 38553 MW; EC614C84C65208CB CRC64;

```

Query Match 20.0%; Score 7; DB 5; Length 341;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 SESSTPS 10
 |||||
 Db 234 SESSTPS 240

OY 9 PSTGFS 15
 |||||
 Db 286 PSTGFS 292
 Search completed: July 24, 2003, 12:25:49
 Job time : 12.7613 secs

RESULT 14

Q8NSR7 PRELIMINARY; PRT; 342 AA.
 AC Q8NSR7;
 DT 01-OCT-2002 (TReMBLrel. 22, Created)
 DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
 DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
 DE Hypothetical protein Cg10601.
 GN Cg10601.
 OS Corynebacterium glutamicum (Brevibacterium flavum).
 OC Bacteria; Actinobacteria; Actinobacteriales;
 OC Corynebacteriaceae; Corynebacterium.
 OX NCBI_TaxID=1718;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
 RA Nakagawa S.;
 RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032."
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP005275; BAB97994.1; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 342 AA; 35227 MW; BCAB356CAA77F36B CRC64;

Query Match 20.0%; Score 7; DB 16; Length 342;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 STPSTG 13
 |||||
 Db 316 STPSTG 322

RESULT 15

Q21298 PRELIMINARY; PRT; 422 AA.
 AC Q21298;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE K07F5.14 protein.
 GN K07F5.14.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodertinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hembry C.;
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 investigating biology."
 RL Science 282:2012-2018(1998).
 DR EMBL; Z70284; CAA94285.4; -
 DR WormPep; K07F5.14; CE29594.
 SQ SEQUENCE 422 AA; 48674 MW; 62F033819A0937A6 CRC64;

Query Match 20.0%; Score 7; DB 5; Length 422;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

PA (KRAT/) KRATOCHVIL J D.
 PA (ROBE/) ROBERTS-RAPP L.
 PA (RUS/) RUSSELL J C.
 PA (STRO/) STROUPE S D.
 PI Billingsel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J,
 PI Granados EN, Hodges SC, Klasse MR, Kratochvil JD, Roberts-Rapp L,
 PI Russell JC, Stroupe SD;
 DR WPI; 2002-665429/71.
 XX
 PS Claim 17; Page 44; 58bp; English.
 CC The present invention relates to a number of prostate-specific sequences
 CC derived from the human PS118 gene. These can be used in the detection,
 CC monitoring and treatment of prostate diseases, particularly prostate
 CC cancer. The present sequence is a PS118 protein fragment of the
 CC invention. The coding sequences of the invention were isolated from a
 CC prostate tissue expressed sequence tag (EST) library.
 XX
 SQ Sequence 35 AA;
 Query Match 100.0%; Score 180; DB 23; Length 35;
 Best Local Similarity 100.0%; Pred. No. 1.2e-18;
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SFQSSSTPTGTGFGSKETPSDDRSQSRHMGES 35
 DB 1 SFQSSSTPTGTGFGSKETPSDDRSQSRHMGES 35
 RESULT 2
 AAM50811
 ID AAM50811 standard; Protein; 35 AA.
 AC AAM50811;
 DT 01-MAY-2002 (first entry)
 DE PS118 prostate marker immunogenic polypeptide.
 XX
 KW PS118; prostate; marker; prostate cancer; tumour; metastasis;
 KW benign prostatic hyperplasia; prostatic intraepithelial neoplasia;
 KW prostaticitis; human; diagnosis; therapy; vaccine; immunogen.
 XX
 OS Homo sapiens.
 XX
 PN US200105758-A1.
 PD 27-DEC-2001.
 PF 23-APR-1998; 98US-0065383.
 PR 23-APR-1997; 97US-0842385.
 PA (BILL/) BILLING-MEDEL P A.
 PA (COHE/) COHEN M.
 PA (COPL/) COPLITTIS T L.
 PA (FRIE/) FRIEDMAN P N.
 PA (GORD/) GORDON J.
 PA (GRAN/) GRANADOS E N.
 PA (HODG/) HODGES S C.
 PA (KLA/) KLAS M R.
 PA (KRAT/) KRATOCHVIL J D.
 PA (ROBE/) ROBERTS-RAPP L.
 PI Billings-Medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J,
 PI Granados EN, Hodges SC, Klasse MR, Kratochvil JD, Roberts-Rapp L,

PI Russell JC, Stroupe SD;
 DR WPI; 2002-187683/24.
 XX
 PS Claim 17; Page 44; 57bp; English.
 CC The present sequence is that of an immunogenic polypeptide
 CC comprising amino acids 283-317 of human prostate-specific PS118
 CC polypeptide (see AAM50809). A PS118 consensus sequence (see
 CC ABA91651) is found at least 12 times more often in prostate than
 CC in non-prostate tissue. PS118 polypeptides, polynucleotides,
 CC antibodies, agonists and inhibitors are useful for detecting,
 CC diagnosing, staging, monitoring, prognosticating, preventing and
 CC treating (including by genetic immunisation), or determining the
 CC predisposition of an individual to, diseases and conditions of the
 CC prostate, such as benign prostatic hyperplasia, prostaticitis,
 CC prostatic intraepithelial neoplasia, prostate cancer, tumours and
 CC metastases. The PS118 polypeptides can be produced by expression
 CC of PS118 polynucleotides in transfected host cells, and
 CC immunogenic peptides are useful for raising PS118-specific
 CC antibodies of diagnostic use. The methods and reagents of the
 CC invention may provide an early means of detecting diseases of the
 CC prostate and may also provide new markers which can differentiate
 CC between the clinically important and unimportant prostate cancers
 CC without the use of surgery.
 XX
 SQ Sequence 35 AA;
 Query Match 100.0%; Score 180; DB 23; Length 35;
 Best Local Similarity 100.0%; Pred. No. 1.2e-18;
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SFQSSSTPTGTGFGSKETPSDDRSQSRHMGES 35
 DB 1 SFQSSSTPTGTGFGSKETPSDDRSQSRHMGES 35
 RESULT 3
 AAM85472
 ID AAM85472 standard; Protein; 518 AA.
 AC AAM85472;
 DT 25-MAR-2003 (updated)
 DT 25-FEB-1999 (first entry)
 DE PS118 protein encoded by consensus sequence.
 XX
 KW EST clone; PS118; prostate tumour tissue; prostatic disease; cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO9848054-A1.
 PD 29-OCT-1998.
 PF 23-APR-1998; 98WO-US08239.
 PR 23-APR-1997; 97US-0842385.
 PA (ABBO) ABBOTT LAB.
 PI Billings-medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J,
 PI Granados EN, Hodges SC, Klasse MR, Kratochvil JD, Robertterapp L,
 PI Russell JC, Stroupe SD;
 DR WPI; 1998-610000/51.
 DR N-PSDB; AAV82812.

XX New P118 nucleic acid and proteins - used for diagnosis and
PT treatment of prostatic disease, especially cancer, and also for drug
PT screening
XX
PS Claim 17; Page 93-94; 117pp; English.
XX
XX The present sequence is encoded by consensus P5118 sequence derived from
CC overlapping EST (expressed sequence tag) clones AAV82803-10. The clones
CC were identified from a cDNA library made from prostate tumour tissue.
CC Recombinant P5118 protein is used to detect P5118-specific antibodies,
CC to raise antibodies for detection of P5118 antigens, to screen for
CC specific binding agents (potential therapeutics, and to isolate specific
CC antibodies from serum. Detection of P5118 protein or nucleic acid, which
CC are prostate related, and altered or elevated in prostatic disease, is
CC used for detection, diagnosis, staging, monitoring and prognosis of
CC prostatic disease, particularly cancer, and to identify subjects at
CC risk.
CC (Updated on 25-MAR-2003 to correct PI field.)
CC
XX Sequence 518 AA;
SQ

Query Match 100.0%; Score 180; DB 19; Length 518;
Best Local Similarity 100.0%; Pred. No. 2.8e-17;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFOSESSTPTGTGFSGKETPSEDDRSQSRHMGES 35
DB 283 SFOSESSTPTGTGFSGKETPSEDDRSQSRHMGES 317

RESULT 4
ID AAO19165 standard; Protein: 518 AA.
XX
AC AAO19165;
XX
DT 27-NOV-2002 (first entry)
XX
DE Human prostate-specific P5118 protein fragment #1.
XX
XX Human; prostate; prostate-specific sequence; prostate cancer; P5118;
KW EST; expressed sequence tag; cytostatic; gene therapy.
XX
OS Homo sapiens.
XX
EN US2002086316-A1.
XX
PD 04-JUL-2002.
XX
PF 26-NOV-2001; 2001US-0991681.
XX
PR 23-APR-1998; 98US-0065383.
PR 23-APR-1997; 97US-0842385.
XX
PA (BILL/) BILLINGEL P A.
PA (COHE/) COHEN M.
PA (COLP/) COLPITTS T L.
PA (FRIE/) FRIEDMAN P N.
PA (GORD/) GORDAN J.
PA (GRAN/) GRANADOS E N.
PA (HODG/) HODGES S C.
PA (KLAS/) KLAAS M R.
PA (KRAT/) KRATOCHVIL J D.
PA (ROBE/) ROBERTS-RAPP L.
PA (RUSSE/) RUSSELL J C.
PA (STRO/) STROUPE S D.
XX
PI Billingsel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;
PI Granados EN, Hodges SC, Klaas MR, Kratochvil JD, Roberts-Rapp L;
PI Russell JC, Stroupe SD;
XX WPI; 2002-665429/71.
DR

XX Novel P5118 polypeptide for detecting, diagnosing, staging, monitoring,
PT prognosticating, preventing, treating, or determining predisposition of
PT individual to diseases and conditions of prostate, e.g. prostate
PT cancer -
XX
PS Claim 17; Page 42-43; 58pp; English.
XX
XX The present invention relates to a number of prostate-specific sequences
CC derived from the human P5118 gene. These can be used in the detection,
CC monitoring and treatment of prostate diseases, particularly prostate
CC cancer. The present sequence is a P5118 protein fragment of the
CC invention. The coding sequences of the invention were isolated from a
CC prostate tissue expressed sequence tag (EST) library.
XX
SQ Sequence 518 AA;
SQ

Query Match 100.0%; Score 180; DB 23; Length 518;
Best Local Similarity 100.0%; Pred. No. 2.8e-17;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFOSESSTPTGTGFSGKETPSEDDRSQSRHMGES 35
DB 283 SFOSESSTPTGTGFSGKETPSEDDRSQSRHMGES 317

RESULT 5
ID AAM50809 standard; Protein: 518 AA.
XX
AC AAM50809;
XX
DT 01-MAY-2002 (first entry)
XX
DE P5118 prostate marker partial sequence.
XX
XX P5118; prostate; marker; prostate cancer; tumour; metastasis;
KW benign prostatic hyperplasia; prostatic intraepithelial neoplasia;
KW prostatic; human; diagnosis; therapy; vaccine.
XX
OS Homo sapiens.
XX
EN US2001055758-A1.
XX
PD 27-DEC-2001.
XX
PF 23-APR-1998; 98US-0065383.
XX
PR 23-APR-1997; 97US-0842385.
XX
PA (BILL/) BILLING-MEDEL P A.
PA (COHE/) COHEN M.
PA (COLP/) COLPITTS T L.
PA (FRIE/) FRIEDMAN P N.
PA (GORD/) GORDON J.
PA (GRAN/) GRANADOS E N.
PA (HODG/) HODGES S C.
PA (KLAS/) KLAAS M R.
PA (KRAT/) KRATOCHVIL J D.
PA (ROBE/) ROBERTS-RAPP L.
XX
PI Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;
PI Granados EN, Hodges SC, Klaas MR, Kratochvil JD, Roberts-Rapp L;
PI Russell JC, Stroupe SD;
XX WPI; 2002-187683/24.
XX N-PSDB; ABA91651.
DR
XX
PT Detecting presence of target P5118 polynucleotide in test sample,
PT useful for detecting, diagnosing, staging, monitoring, prognosticating,
PT preventing or treating or determining predisposition to prostate
PT disease -
XX

PS Claim 17; Page 42-43; 57bp; English.

CC The present sequence is that of a human prostate-specific PS18
 CC polypeptide, as predicted from a partial consensus cDNA sequence
 CC (see ABA91651), and lacking the N-terminal region. The PS18
 CC consensus sequence is found at least 12 times more often in
 CC prostate than in non-prostate tissue. PS18 polypeptides,
 CC including derivatives of the present sequence, polynucleotides,
 CC antibodies, agonists and inhibitors are useful for detecting,
 CC diagnosing, staging, monitoring, prognosticating, preventing and
 CC treating (including by genetic immunisation), or determining the
 CC predisposition of an individual to, diseases and conditions of the
 CC prostate, such as benign prostatic hyperplasia, prostatitis,
 CC prostatic intraepithelial neoplasia, prostate cancer, tumours and
 CC metastases. The PS18 polypeptides can be produced by expression
 CC of PS18 polynucleotides in transfected host cells. The methods
 CC and reagents of the invention may provide an early means of
 CC detecting diseases of the prostate and may also provide new markers
 CC which can differentiate between the clinically important and
 CC unimportant prostate cancers without the use of surgery.

SQ Sequence 518 AA;

Query Match 100.0%; Score 180; DB 23; Length 518;
 Best Local Similarity 100.0%; Pred. No. 2.8e-17;
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFOSSSSTPSTGFGSGKETPSEDPRSOSREHMGES 35
 |||
 DB 283 SFOSSSSTPSTGFGSGKETPSEDPRSOSREHMGES 317

RESULT 6
 ID ABA909728 standard; Protein; 1807 AA.

AC ABA909728;
 XX
 DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #9719.

KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KM food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.
 XX
 XX WO200175067-A2.
 XX
 XX 11-OCT-2001.
 XX
 XX 30-MAR-2001; 2001WO-US08631.
 XX
 XX 31-MAR-2000; 2000US-0540217.
 XX
 XX 23-AUG-2000; 2000US-0649167.
 XX
 XX (HYSE-) HYSEQ INC.
 XX
 XX Drmanac RT, Liu C, Tang YT;
 XX
 XX WPI: 2001-639362/73.
 XX
 XX N-PSDB; AAS73915.

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX
 XX Claim 20; SEQ ID No 40087; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABA90010-ABA93077 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WFO
 CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 1807 AA;

Query Match 100.0%; Score 180; DB 22; Length 1807;
 Best Local Similarity 100.0%; Pred. No. 1.2e-16;
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFOSSSSTPSTGFGSGKETPSEDPRSOSREHMGES 35
 |||
 DB 1572 SFOSSSSTPSTGFGSGKETPSEDPRSOSREHMGES 1606

RESULT 7
 ID ABA94835 standard; Protein; 1839 AA.

AC ABA94835;
 XX
 DT 25-FEB-2003 (first entry)

DE Human protein SEQ ID 495.

KM Human; expressed sequence tag; EST;
 KM haematopoietic disorder; central nervous system disease; viral infection;
 KM peripheral nervous system disease; non-healing wound; infectious disease;
 KM immune deficiency; immune disorder; bacterial infection; allergy; cancer;
 KM fungal infection; autoimmune disorder; coagulation disorder; neutrotic;
 KM antiallergic; antiinflammatory; immunosuppressive; neuroprotective;
 KM cyostatic; haemostatic; viroicide; antibacterial; fungicide;
 KM immunostimulant; cerebroprotective.

OS Homo sapiens.
 XX
 XX WO200259260-A2.
 XX
 XX 01-AUG-2002.
 XX
 XX 16-NOV-2001; 2001WO-US42950.
 XX
 XX 17-NOV-2000; 2000US-0714936.
 XX
 XX (HYSE-) HYSEQ INC.
 XX
 XX Tang YT, Goodrich RW, Liu C, Zhou P, Adundi V, Zhang J, Zhao QA;
 XX
 XX Ren F, Xue AJ, Yang Y, Wehrman T, Drmanac RT;
 XX
 XX WPI: 2002-590824/63.
 XX
 XX N-PSDB; ABA99421.

PT New isolated polynucleotide, useful in research, diagnostic or
 PT therapeutic methods, e.g. preventing or treating disorders involving
 PT aberrant protein expression or biological activity -
 XX
 XX Claim 20; SEQ ID 495; 394pp; English.

CC The present invention relates to novel human coding sequences
 CC (AB093268-AB093608) and proteins (ABP64682-ABP65022). The sequences are
 CC useful in therapeutic, diagnostic and research methods. The
 CC polynucleotides may be used in the field of molecular biology as
 CC hybridisation probes, primers for PCR, for chromosome and gene mapping,
 CC for the recombinant production of protein, or in generation of anti-sense
 CC DNA or RNA. The polynucleotides are useful in diagnostics as expressed
 CC sequence tags (ESTs) for identifying expressed genes or for physical
 CC mapping of the human genome. The proteins may be used as molecular weight
 CC markers, or as nutritional sources or supplements. The proteins may be
 CC used to maintain and expand cell population in a totipotent or
 CC pluripotent state useful for re-engineering damaged or diseased
 CC tissues, transplantation, manufacture of bio-pharmaceuticals or the
 CC development of bio-sensors. The polynucleotides and proteins are useful
 CC for preventing, treating or ameliorating disorders involving aberrant
 CC protein expression or biological activity, e.g. hematopoietic disorders,
 CC central/peripheral nervous system diseases, mechanical and traumatic
 CC disorders, non-healing wounds, immune deficiencies and disorders,
 CC infectious diseases caused by viral, bacterial or fungal infection,
 CC autoimmune disorders, allergic reactions and conditions, coagulation
 CC disorders, or cancer. The polynucleotide sequences of the invention were
 CC assembled from ESTs isolated mainly by sequencing by hybridisation, and
 CC in some cases, sequences obtained from one or more public databases.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1839 AA;

Query Match 100.0%; Score 180; DB 23; Length 1839;
 Best Local Similarity 100.0%; Pred. No. 1.2e-16;
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SFOSESSTPTGCGSKETPSEDDRSOSREHMGES 35
 Db 1604 SFOSESSTPTGCGSKETPSEDDRSOSREHMGES 1638

RESULT 8

ABG09731 ID ABG09731 standard; Protein; 1982 AA.

XX ABG09731;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #9722.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

OS WO200175067-A2.

PN 11-OCT-2001.

PD 30-MAR-2001; 2001WO-US08631.

PF 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

XX N-PSDB; AAS73918.

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -

XX Claim 20; SEQ ID NO 40090; 103pp; English.

PS The invention relates to isolated polynucleotide (I) and
 XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polynucleotide chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG03077 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1982 AA;

Query Match 100.0%; Score 180; DB 22; Length 1982;
 Best Local Similarity 100.0%; Pred. No. 1.3e-16;
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SFOSESSTPTGCGSKETPSEDDRSOSREHMGES 35
 Db 1686 SFOSESSTPTGCGSKETPSEDDRSOSREHMGES 1720

RESULT 9

ABBS9318 ID ABBS9318 standard; Protein; 1061 AA.

XX ABBS9318;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 4746.

XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.

XX Drosophila melanogaster.

OS WO200171042-A2.

PN 27-SEP-2001.

PD 23-MAR-2001; 2001WO-US09231.

PF 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

DR WPI; 2001-656860/75.

XX N-PSDB; ABL03421.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -

PS Disclosure; SEQ ID NO 4746; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent.
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins
 CC (AB57737-AB872072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX Sequence 1061 AA;
 SQ
 Query Match 33.9%; Score 61; DB 22; Length 1061;
 Best Local Similarity 51.9%; Pred. No. 8.7;
 Matches 14; Conservative 4; Mismatches 9; Indels 0; Gaps 0;
 QY 4 SSSSTPSTGSGSKETPSDDRSQSRRE 30
 Db 39 SKSRPSSAGVVIDETQSEEESSQSS 65
 RESULT 10
 ABG06957
 ID ABG06957 standard; Protein; 253 AA.
 AC ABG06957;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #6948.
 XX
 KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KM food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI, 2001-639362/73.
 DR N-PSDB; AAS71144.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX
 PS Claim 20; SEQ ID No 37316; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (II) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.

CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX Sequence 253 AA;
 SQ
 Query Match 32.8%; Score 59; DB 22; Length 253;
 Best Local Similarity 42.9%; Pred. No. 3.2;
 Matches 15; Conservative 5; Mismatches 11; Indels 4; Gaps 2;
 QY 4 SSSSTPST---GFGSGK-ETPSDDRSQSRHMG 34
 Db 122 SATSTPSSYLPSGLGDKSETPSDEEBEBEBEBEBE 156
 RESULT 11
 ABB60540
 ID ABB60540 standard; Protein; 1403 AA.
 AC ABB60540;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 8412.
 XX
 KM Drosophila; developmental biology; cell signalling; insecticide;
 KM pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI, 2001-6556860/75.
 DR N-PSDB; ABL04643.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure; SEQ ID NO 8412; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins
 CC (AB57737-AB872072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX Sequence 1403 AA;
 SQ
 Query Match 32.8%; Score 59; DB 22; Length 1403;

Best Local Similarity 38.1%; Pred. No. 23;
Matches 16; Conservative 5; Mismatches 13; Indels 8; Gaps 1;
QY 1 SFSSTPTGSGFSGKET-----PSEDDRSQSRHMG 34
Db 1104 SSSSSSTSSLSLAGESSDFNELALAEKELSSQSGTHMG 1145

RESULT 12
AAU42806
ID AAU42806 standard; Protein; 73 AA.
XX
AC AAU42806;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #3702.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PE 20-APR-2001; 2001WO-US12865.
XX
PR 21-APR-2000; 2000US-199047P.
XX
PR 02-JUN-2000; 2000US-208841P.
XX
PR 07-JUL-2000; 2000US-216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI: 2001-616774/71.
DR N-PSDB; AAS59518.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
PS Example 1; SEQ ID NO 4001; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 73 AA;
Query Match 32.2%; Score 58; DB 22; Length 73;

Best Local Similarity 34.3%; Pred. No. 1;
Matches 12; Conservative 6; Mismatches 15; Indels 2; Gaps 1;
QY 3 QSESTPTGSGFSGKETPSEDDR--SOSRHHMGS 35
Db 12 KSEQSKPGAGVGEGQRRQDDAKPGHGRDEVGA 46

RESULT 13
ABB70235
ID ABB70235 standard; Protein; 345 AA.
XX
AC ABB70235;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 37497.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PE 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li FWD, Myers EW;
XX
DR WPI: 2001-656860/75.
DR N-PSDB; ABL14338.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 37497; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 345 AA;
Query Match 31.7%; Score 57; DB 22; Length 345;
Best Local Similarity 38.7%; Pred. No. 8.9;
Matches 12; Conservative 5; Mismatches 14; Indels 0; Gaps 0;
QY 5 ESSTPTGSGFSGKETPSEDDRSQSRHMG 35
Db 123 KESTYNSGFAPPTQVPTDVKVDSGSHMDTS 153
RESULT 14
ABR40769
ID ABR40769 standard; Protein; 530 AA.
XX
AC ABR40769;

```

XX 16-MAY-2003 (first entry)
DT Glycine max oil trait related protein sequence SEQ ID NO:358.
XX
XX Glycine max oil trait related protein sequence SEQ ID NO:358.
DE
XX Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;
XX receptor-like protein kinase; mitogen activated protein kinase; oil;
XX LIP5-like transcription factor caleosin; ATP citrate lyase; SNF1;
XX CKC-like transcription factor; antisense inhibition; co-suppression;
XX transgenic plant.
XX
XX Glycine max.
OS
XX WO2003002751-A2.
XX
XX 09-JAN-2003.
XX
XX 27-JUN-2002; 2002WO-US20152.
XX
XX 29-JUN-2001; 2001US-301913P.
XX
XX (DUPO) DU PONT DE NEMOURS & CO E I.
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Allen SM, Allen WB, Cahoon RS, Epelbaum S, Farnedu OO, Harvell LT;
XX Jones TJ, Kinney AJ, Klein TM, Li C, Oliveira IC, Sakai H;
XX Shen B, Tarczy-Hartmann MC;
XX WPI; 2003-201509/19.
XX N-PSDB; ACC00804.
XX
XX Novel nucleotide fragment encoding polypeptides having receptor-like
XX protein kinase activity, caleosin-like activity, useful for altering
XX oil phenotypes in plants such as sunflower, coconut, soybean, wheat and
XX rice.
XX
XX Claim 12; Page 377-379; 542pp; English.
XX
XX The present invention describes an isolated nucleotide fragment (I)
XX comprising a nucleic acid sequence (NS) chosen from a NS encoding a
XX polypeptide (PP) having receptor-like protein kinase activity, mitogen
XX activated protein (MAP)-kinase activity, lipid-like transcription factor
XX activity, caleosin-like activity, ATP citrate lyase activity, SNF1-like
XX activity and CKC-like transcription factor activity. Also described:
XX (1) complement (II) of (I); (2) a chimeric construct (III) comprising (I)
XX or (II), operably linked to a regulatory sequence; (3) a plant (IV)
XX comprising (III) in its genome; (4) seeds (V) obtained from (IV); and
XX (5) oil obtained from (V). (I) or its part can be used in antisense
XX inhibition or co-suppression in a transformed plant. (III) is useful for
XX altering the oil phenotype in a plant such as corn, soybean, wheat, rice,
XX canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for
XX creating transgenic plants having altered lipid profiles. (I) can also
XX be used as a hybridization probe. ACC00626 to ACC00868 and ABR40591 to
XX ABR40879 represent sequences used in the exemplification of the present
XX invention.
XX
XX Sequence 530 AA;
SQ

```

```

Query Match 31.7%; Score 57; DB 24; Length 530;
Best Local Similarity 46.2%; Pred. No. 15;
Matches 12; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

```

```

QY 6 SSTPTGSGSKETPSDDRSQSRHM 31
DB 322 NSTLPIGLSGKXKNTSDASSEKSH 347

```

```

RESULT 15
ABG50041
ID ABG50041 standard; Peptide; 102 AA.
XX
XX AC ABG50041;
XX

```

```

DT 25-FEB-2003 (first entry)
XX
XX Human liver peptide, SEQ ID No 28689.
XX
XX Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
XX hypercholesterolaemia; coronary heart disease.
XX
XX Homo sapiens.
XX
XX WO200157273-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00664.
XX
XX 04-FEB-2000; 2000US-0180312.
XX
XX 26-MAY-2000; 2000US-0207456.
XX
XX 30-JUN-2000; 2000US-0608408.
XX
XX 03-AUG-2000; 2000US-0632366.
XX
XX 21-SEP-2000; 2000US-0234687.
XX
XX 27-SEP-2000; 2000US-0236359.
XX
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488898/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analysing gene expression in human adult liver.
XX
XX Claim 27; SEQ ID No 28689; 658pp; English.
XX
XX The invention relates to a single exon nucleic acid probe (SENP) (I) for
XX measuring human gene expression in a sample derived from human adult
XX liver, comprising one of 13109 defined nucleotide sequences given in the
XX specification (or complements/ fragments). The probe hybridises at high
XX stringency to a nucleic acid molecule expressed in the human adult
XX liver. (I) may be used for predicting, measuring and displaying gene
XX expression in samples derived from human adult liver. The genes
XX identified may be involved in genetic liver diseases such as cirrhosis,
XX hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which
XX is associated with coronary heart disease. ABG47348-ABG59930 represent
XX human liver single exon encoded peptides of the invention.
XX Note: The sequence information for this patent does not appear in the
XX printed specification but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 102 AA;
SQ

```

```

Query Match 31.1%; Score 56; DB 22; Length 102;
Best Local Similarity 40.0%; Pred. No. 3;
Matches 12; Conservative 6; Mismatches 8; Indels 4; Gaps 1;

```

```

QY 3 QSESTPTGSGSKETPSDDRSQSRHM 32
DB 27 QNESQSPQ---EPBEGPSDDKAGBEBM 52

```

```

Search completed: July 24, 2003, 11:53:40
Job time: 14.3763 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 11:53:53 ; Search time 3.12592 Seconds
(without alignments)
473.743 Million cell updates/sec

Title: US-09-991-681-29

Perfect score: 180
Sequence: 1 SFQSESTPSTGFGSKETPSDDRSQSREHMGES 35

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep: *
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep: *
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep: *
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep: *
5: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pep: *
6: /cgn2_6/ptodata/1/1aa/backfile1.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	180	100.0	35	US-09-065-383-29	Sequence 29, Appl
2	180	100.0	518	US-09-065-383-27	Sequence 27, Appl
3	53.5	29.7	737	US-09-772-647-4	Sequence 4, Appl
4	52.5	29.2	717	US-08-910-925-1	Sequence 1, Appl
5	50	27.8	309	US-08-465-167A-24	Sequence 24, Appl
6	50	27.8	309	US-08-993-118-10	Sequence 10, Appl
7	50	27.8	309	US-08-845-528C-10	Sequence 10, Appl
8	50	27.8	309	US-08-627-820-24	Sequence 10, Appl
9	50	27.8	309	US-09-066-281B-10	Sequence 10, Appl
10	50	27.8	333	US-09-107-532A-6475	Sequence 10, Appl
11	50	27.8	401	US-08-549-004A-5	Sequence 5, Appl
12	50	27.8	401	US-09-051-982A-5	Sequence 5, Appl
13	50	27.8	1261	US-09-208-742-4	Sequence 4, Appl
14	50	27.8	1261	US-09-332-295-2	Sequence 2, Appl
15	50	27.8	1261	US-09-709-979-2	Sequence 2, Appl
16	49.5	27.5	99	US-09-216-393B-20	Sequence 20, Appl
17	49.5	27.5	674	US-08-893-852A-1	Sequence 1, Appl
18	49.5	27.5	703	US-08-910-925-4	Sequence 4, Appl
19	49	27.2	683	US-09-620-412C-357	Sequence 35, App
20	49	27.2	683	US-09-598-419-357	Sequence 35, App
21	49	27.2	821	US-09-252-991A-30781	Sequence 30781, A
22	49	27.2	821	US-09-556-877-195	Sequence 195, App
23	49	27.2	821	US-09-620-412C-195	Sequence 195, App
24	49	27.2	821	US-09-598-419-195	Sequence 195, App
25	49	27.2	1776	US-09-556-877-179	Sequence 179, App
26	49	27.2	1776	US-09-620-412C-179	Sequence 179, App
27	49	27.2	1776	US-09-598-419-179	Sequence 179, App

28	48.5	26.9	335	4	US-09-106-872A-17	Sequence 17, Appl
29	48.5	26.9	568	1	US-08-320-559-30	Sequence 30, Appl
30	48.5	26.9	568	3	US-08-545-860D-30	Sequence 30, Appl
31	48.5	26.9	568	5	PCT-US94-004496-30	Sequence 30, Appl
32	48	26.7	432	4	US-09-252-991A-31253	Sequence 31253, A
33	48	26.7	714	4	US-09-347-878-16	Sequence 16, Appl
34	47.5	26.4	590	3	US-08-893-852A-4	Sequence 4, Appl
35	47.5	26.4	590	3	US-08-821-818-2	Sequence 2, Appl
36	47.5	26.4	590	4	US-09-052-753B-2	Sequence 2, Appl
37	47	26.1	300	4	US-09-277-078-2	Sequence 2, Appl
38	47	26.1	365	2	US-08-481-814A-10	Sequence 10, Appl
39	47	26.1	1704	3	US-08-485-355B-40	Sequence 40, Appl
40	47	26.1	2265	2	US-08-149-097D-36	Sequence 36, Appl
41	47	26.1	2509	2	US-08-149-097D-35	Sequence 35, Appl
42	46.5	25.8	98	2	US-08-481-658B-50	Sequence 50, Appl
43	46.5	25.8	98	2	US-08-477-504A-50	Sequence 50, Appl
44	46.5	25.8	98	2	US-08-466-756A-50	Sequence 50, Appl
45	46.5	25.8	98	2	US-08-485-862B-50	Sequence 50, Appl

ALIGNMENTS

RESULT 1
US-09-065-383-29
Sequence 29, Application US/09065383
Patent No. 6391543
GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GORDON, JULIAN
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KLASS, MICHAEL R.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: ROBERTS-RAFF, LISA
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/065,383
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/842,385
FILING DATE: 23-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6084.US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 35 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6391543e
US-09-065-383-29

Query Match 100.0%; Score 180; DB 4; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.3e-18;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SF0SESTPSTGFGSGKETPSEDDRSQSRHMGES 35
Db 1 SF0SESTPSTGFGSGKETPSEDDRSQSRHMGES 35

RESULT 2
US-09-065-383-27
Sequence 27, Application US/09065383
Patent No. 6391543

GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GORDON, JULIAN
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KLAAS, MICHAEL R.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: ROBERTS-RAPP, LISA
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
TITLE OF INVENTION: FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/065,383
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/842,385
FILING DATE: 23-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6084.US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 518 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6391543e
US-09-065-383-27

'Query Match 100.0%; Score 180; DB 4; Length 518;
Best Local Similarity 100.0%; Pred. No. 3.3e-17;

Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SF0SESTPSTGFGSGKETPSEDDRSQSRHMGES 35
Db 283 SF0SESTPSTGFGSGKETPSEDDRSQSRHMGES 317

RESULT 3
US-09-772-647-4
Sequence 4, Application US/09772647
Patent No. 6521815

GENERAL INFORMATION:
APPLICANT: Verna, Ajit K.
APPLICANT: Reddig, Peter U.
APPLICANT: Jansen, Aaron P.
TITLE OF INVENTION: Animal Model System for Squamous Cell Carcinoma
FILE REFERENCE: 960296.97613
CURRENT APPLICATION NUMBER: US/09/772,647
CURRENT FILING DATE: 2001-01-30
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 737
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: T7 tag and
OTHER INFORMATION: mouse protein kinase C epsilon coding sequence
US-09-772-647-4

Query Match 29.7%; Score 53.5; DB 4; Length 737;
Best Local Similarity 48.0%; Pred. No. 26;
Matches 12; Conservative 5; Mismatches 3; Indels 5; Gaps 1;

Qy 4 SEESTPSTGFGSGKETPSEDDRSQS 28
Db 327 AESQPASG-----NSPSEDDRSKS 346

RESULT 4
US-08-910-925-1
Sequence 1, Application US/08910925
Patent No. 6162601

GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
APPLICANT: Shah, Purvi
TITLE OF INVENTION: HUMAN PININ SPLICE VARIANT
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/910,925
FILING DATE: Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0365 US
TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 717 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: FIBROTO1
CLONE: 53219
US-08-910-925-1

Query Match 29.2%; Score 52.5; DB 3; Length 717;
Best Local Similarity 41.2%; Pred. No. 34;
Matches 14; Conservative 6; Mismatches 9; Indels 5; Gaps 1;

OY 1 SPOSESTPTGSGKGTPESEDRSQSR 29
DB 603 SRSSTSSSTGSGSSRDSSTSSSTSSSR 636

RESULT 5
US-08-465-167A-24
Sequence 24, Application US/08465167A
Patent No. 5750395
GENERAL INFORMATION:
APPLICANT: Fikes, John D.
APPLICANT: Livingston, Brian D.
APPLICANT: Sette, Alessandro D.
APPLICANT: Sidney, John C.
TITLE OF INVENTION: DNA ENCODING MAGE-1 C-TERMINAL
TITLE OF INVENTION: IMMUNOGENIC PEPTIDES (as amended)
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,167A
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/103,623
FILING DATE: 06-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: Parmelee, Steven W.
REGISTRATION NUMBER: 31,990
REFERENCE/DOCKET NUMBER: 14137-60-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-467-9600
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 309 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-465-167A-24

Query Match 27.8%; Score 50; DB 1; Length 309;
Best Local Similarity 42.4%; Pred. No. 28;
Matches 14; Conservative 4; Mismatches 13; Indels 2; Gaps 1;

OY 3 QSESSTPTGSGKGTPESEDRSQSRHMGES 35
DB 59 QGASAPPTTINFTRQRPSEG--SSSRBEGSPS 89

RESULT 6
US-08-993-118-10
Sequence 10, Application US/08993118
Patent No. 5997872
GENERAL INFORMATION:
APPLICANT: LUCAS, Sophie;
APPLICANT: DE SMET, Charles;
APPLICANT: BOON-FALLEUR, Thierry
TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE CODING FOR TUMOR
TITLE OF INVENTION: REJECTION ANTIGEN PRECURSOR MAGE-C1 AND USES
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/993,118
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/845,528
FILING DATE: April 25, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mary Anne Schofield
REGISTRATION NUMBER: 36,669
REFERENCE/DOCKET NUMBER: LUD 5455
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 309
TYPE: amino acids
STRANDEDNESS: single stranded
TOPOLOGY: linear
US-08-993-118-10

Query Match 27.8%; Score 50; DB 2; Length 309;
Best Local Similarity 42.4%; Pred. No. 28;
Matches 14; Conservative 4; Mismatches 13; Indels 2; Gaps 1;

OY 3 QSESSTPTGSGKGTPESEDRSQSRHMGES 35
DB 59 QGASAPPTTINFTRQRPSEG--SSSRBEGSPS 89

RESULT 7
US-08-845-528C-10
Sequence 10, Application US/08845528C
Patent No. 6027924
GENERAL INFORMATION:
APPLICANT: LUCAS, Sophie;
APPLICANT: DE SMET, Charles;
APPLICANT: BOON-FALLEUR, Thierry
TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE CODING FOR TUMOR
TITLE OF INVENTION: REJECTION ANTIGEN PRECURSOR MAGE-C1 AND USES
NUMBER OF SEQUENCES: 14

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/845,528C
FILING DATE: April 25, 1997
CLASSIFICATION: 4335
ATTORNEY/AGENT INFORMATION:
NAME: Mary Anne Schofield
REGISTRATION NUMBER: 36,669
REFERENCE/DOCKET NUMBER: LUD 5455
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 309
TYPE: amino acids
STRANDEDNESS: single stranded
TOPOLOGY: linear

```

```

?      REGISTRATION NUMBER: 31,990
?      REFERENCE/DOCKET NUMBER: 14137-60
?      TELECOMMUNICATION INFORMATION:
?      TELEPHONE: (206) 467-9600
?      TELEFAX: (415) 543-5043
?      INFORMATION FOR SEQ ID NO: 24:
?      SEQUENCE CHARACTERISTICS:
?      LENGTH: 309 amino acids
?      TYPE: amino acid
?      STRANDEDNESS: single
?      TOPOLOGY: linear
?      MOLECULE TYPE: protein
?      SEQUENCE DESCRIPTION: SEQ ID NO: 24:
?      JS-08-627-820-24

```


TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 401 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: linear
US-09-051-982A-5

Query Match 27.8%; Score 50; DB 3; Length 401;
Best Local Similarity 42.4%; Pred. No. 38;
Matches 14; Conservative 3; Mismatches 14; Indels 2; Gaps 1;

Qy 1 SPTSTGTFSGKETPSDDRSQSRHMGES 33
Db 182 SVNGNRRTYSSGSGSPHPS--SRSSSRHNG 212

RESULT 13
US-09-208-742-4
; Sequence 4, Application US/09208742
; Patent No. 6174679
; GENERAL INFORMATION:

; APPLICANT: Kaufmann, Joerg
; TITLE OF INVENTION: C1F150/HTAF1150 is Necessary for Cell
; FILE REFERENCE: 1453.002
; CURRENT APPLICATION NUMBER: US/09/208,742
; CURRENT FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1261
; TYPE: PRT
; ORGANISM: human
US-09-208-742-4

Query Match 27.8%; Score 50; DB 3; Length 1261;
Best Local Similarity 36.7%; Pred. No. 1.5e+02;
Matches 11; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

Qy 6 SPTSTGTFSGKETPSDDRSQSRHMGES 35
Db 840 NSREGTGSNGKRRERYTENRGSSRSHSGT 869

RESULT 14
US-09-332-295-2
; Sequence 2, Application US/09332295
; Patent No. 6303372
; GENERAL INFORMATION:
; APPLICANT: Kaufmann, Joerg
; TITLE OF INVENTION: C1F130 INHIBITS CELL CYCLE PROGRESSION
; FILE REFERENCE: 200130.456 / 1513.003
; CURRENT APPLICATION NUMBER: US/09/332,295
; CURRENT FILING DATE: 1999-06-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1261
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-332-295-2

Query Match 27.8%; Score 50; DB 4; Length 1261;
Best Local Similarity 36.7%; Pred. No. 1.5e+02;
Matches 11; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

Qy 6 SPTSTGTFSGKETPSDDRSQSRHMGES 35
Db 840 NSREGTGSNGKRRERYTENRGSSRSHSGT 869

RESULT 15
US-09-709-979-2
; Sequence 2, Application US/09709979
; Patent No. 6423822
; GENERAL INFORMATION:

; APPLICANT: Kaufmann, Joerg
; TITLE OF INVENTION: C1F130 INHIBITS CELL CYCLE PROGRESSION
; FILE REFERENCE: 200130.456 / 1513.003
; CURRENT APPLICATION NUMBER: US/09/709,979
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 09/332,295
; PRIOR FILING DATE: 1999-06-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1261
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-709-979-2

Query Match 27.8%; Score 50; DB 4; Length 1261;
Best Local Similarity 36.7%; Pred. No. 1.5e+02;
Matches 11; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

Qy 6 SPTSTGTFSGKETPSDDRSQSRHMGES 35
Db 840 NSREGTGSNGKRRERYTENRGSSRSHSGT 869

Search completed: July 24, 2003, 12:16:23
Job time : 4.12592 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: July 24, 2003, 11:57:13 ; Search time 4.04832 Seconds
(without alignments)
1026.746 Million cell updates/sec

Title: US-09-991-681-29

Perfect score: 180
Sequence: 1 SFOSSSTPTGFGSKETPEEDRSQSREHMGES 35

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	56	31.1	102	9	US-09-864-761-35901
2	53.5	29.7	737	15	US-10-228-931-4
3	52.5	29.2	91	14	US-10-090-035-16
4	51	28.3	42	10	US-09-955-807-9
5	51	28.3	65	10	US-09-955-807-5
6	51	28.3	65	10	US-09-955-807-10
7	51	28.3	77	10	US-09-955-807-4
8	51	28.3	79	10	US-09-955-807-3
9	51	28.3	80	10	US-09-955-807-16
10	51	28.3	89	10	US-09-955-807-15
11	51	28.3	105	10	US-09-955-807-2
12	51	28.3	105	10	US-09-955-807-14
13	51	28.3	758	9	US-09-904-987-5
14	50	27.8	212	11	US-09-764-891-4686
15	50	27.8	253	9	US-09-864-761-37733

16	50	27.8	275	9	US-09-755-456-9	Sequence 9, Appli
17	50	27.8	309	10	US-09-766-889A-2	Sequence 2, Appli
18	50	27.8	309	14	US-10-085-108-10	Sequence 10, Appli
19	50	27.8	413	9	US-09-801-574-2	Sequence 2, Appli
20	50	27.8	428	15	US-10-156-761-11846	Sequence 11846, A
21	50	27.8	856	10	US-09-738-626-3515	Sequence 3515, Ap
22	50	27.8	1031	9	US-09-815-242-10932	Sequence 10932, A
23	50	27.8	1261	14	US-10-147-268-2	Sequence 2, Appli
24	50	27.8	1261	15	US-10-338-279-2	Sequence 444, App
25	50	27.8	1770	9	US-09-841-132-444	Sequence 20, Appli
26	49.5	27.5	99	9	US-09-216-393-20	Sequence 12867, A
27	49.5	27.5	232	15	US-10-156-761-12867	Sequence 384, App
28	49.5	27.5	515	10	US-09-801-368-384	Sequence 195, App
29	49.5	27.5	737	10	US-09-771-161A-195	Sequence 3475, Ap
30	49.5	27.5	1908	15	US-10-128-714-3475	Sequence 8475, Ap
31	49.5	27.5	2022	15	US-10-128-714-8475	Sequence 5912, Ap
32	49	27.2	331	9	US-09-815-242-5912	Sequence 12872, A
33	49	27.2	331	9	US-09-815-242-12872	Sequence 13146, A
34	49	27.2	331	9	US-09-815-242-13146	Sequence 350, App
35	49	27.2	489	9	US-09-876-889-350	Sequence 195, App
36	49	27.2	683	9	US-09-841-132-357	Sequence 179, App
37	49	27.2	821	9	US-09-841-132-195	Sequence 56, Appli
38	49	27.2	1776	9	US-09-841-132-179	Sequence 14105, A
39	49	27.2	2011	15	US-10-176-847-56	Sequence 1053, Ap
40	48.5	26.9	550	15	US-10-156-761-14105	Sequence 4135, Ap
41	48.5	26.9	724	11	US-09-925-300-1053	Sequence 9761, Ap
42	48	26.7	121	11	US-09-764-891-4135	Sequence 6577, Ap
43	48	26.7	324	15	US-10-156-761-9761	Sequence 85, Appli
44	48	26.7	324	10	US-09-738-626-6577	
45	48	26.7	699	15	US-10-198-070-85	

ALIGNMENTS

RESULT 1
US-09-864-761-35901
Sequence 35901, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OR INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecm1ca-X-1
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US/09/864,761
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30

```
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 35901
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AI096678.8
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.96
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.98
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
; OTHER INFORMATION: EST_HUMAN HIT: BE792924.1, EVALUATE 1.40e-01
US-09-864-761-35901
```

```
Query Match          31.1%; Score 56; DB 9; Length 102;
Best Local Similarity 40.0%; Pred. No. 6.4;
Matches 12; Conservative 6; Mismatches 8; Indels 4; Gaps 1;

Qy 3 QSESTPTGTGFGSGKETPSDDRSQSRHM 32
Db 27 QNESQSPQ----EPREGPSEDDKAREGEHM 52
```

```
RESULT 2
US-10-228-931-4
; Sequence 4, Application US/10228931
; Publication No. US20030051258A1
; GENERAL INFORMATION:
; APPLICANT: Verma, Ajit K
; APPLICANT: Reddig, Peter J
; APPLICANT: Jansen, Aaron P
; TITLE OF INVENTION: Animal Model System for Squamous Cell Carcinoma
; FILE REFERENCE: 960296, 97613
; CURRENT APPLICATION NUMBER: US/10/228,931
; CURRENT FILING DATE: 2002-08-27
; PRIOR APPLICATION NUMBER: US/09/772,647
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 737
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: T7 tag and
; OTHER INFORMATION: mouse protein kinase C epsilon coding sequence
US-10-228-931-4
```

```
Query Match          29.7%; Score 53.5; DB 15; Length 737;
Best Local Similarity 48.0%; Pred. No. 1.2e+02;
Matches 12; Conservative 5; Mismatches 3; Indels 5; Gaps 1;

Qy 4 SSSSTPTGTGFGSGKETPSDDRSQSR 28
Db 327 ASSQPASG-----NSPSEDDRSKS 346
```

```
RESULT 3
US-10-090-035-16
; Sequence 16, Application US/10090035
; Publication No. US20020170089A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; TITLE OF INVENTION: Nucleic Acids Encoding Defense Inducible
; FILE REFERENCE: 35718/24290
; CURRENT APPLICATION NUMBER: US/10/090,035
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/272,227
; PRIOR FILING DATE: 02/28/2001
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-090-035-16

Query Match          29.2%; Score 52.5; DB 14; Length 91;
Best Local Similarity 43.9%; Pred. No. 16;
Matches 18; Conservative 1; Mismatches 11; Indels 11; Gaps 2;

Qy 1 SFQSESTPTGTGFGSG-----KETPSDDRSQSRHM 31
Db 11 SEEVRSVTP-TGFGFGRGVQQCHVVKETPQEIIDRSQSGRH 50
```

```
RESULT 4
US-09-955-807-9
; Sequence 9, Application US/09955807
; Patent No. US20020132996A1
; GENERAL INFORMATION:
; APPLICANT: Lok, Si
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Kindsvogel, Wayne
; APPLICANT: Bort, Susan J.
; TITLE OF INVENTION: Secretory Protein-48
; FILE REFERENCE: 98-17C1
; CURRENT APPLICATION NUMBER: US/09/955,807
; CURRENT FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/102,679
; PRIOR FILING DATE: 1998-10-01
; PRIOR APPLICATION NUMBER: 09/410,603
; PRIOR FILING DATE: 1999-10-01
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-955-807-9
```

```
Query Match          28.3%; Score 51; DB 10; Length 42;
Best Local Similarity 52.2%; Pred. No. 10;
Matches 12; Conservative 3; Mismatches 4; Indels 4; Gaps 1;

Qy 9 PSTGFGSG-----KETPSDDRSQ 27
Db 17 PAVTGTGDSGAKETVSDKRSQ 39
```

```
RESULT 5
US-09-955-807-5
; Sequence 5, Application US/09955807
; Patent No. US20020132996A1
; GENERAL INFORMATION:
; APPLICANT: Lok, Si
; APPLICANT: Sheppard, Paul O.
```



```

1  APPLICANT: Kindsvogel, Wayne
2  APPLICANT: Bott, Susan J.
3  TITLE OF INVENTION: Secretary Protein-48
4  FILE REFERENCE: 98-17C1
5  CURRENT APPLICATION NUMBER: US/09/955,807
6  CURRENT FILING DATE: 2001-09-19
7  PRIOR APPLICATION NUMBER: 60/102,679
8  PRIOR FILING DATE: 1998-10-01
9  PRIOR APPLICATION NUMBER: 09/410,603
10 PRIOR FILING DATE: 1999-10-01
11 NUMBER OF SEQ ID NOS: 17
12 SOFTWARE: FastSeq for Windows Version 3.0
13 SEQ ID NO 5
14 LENGTH: 65
15 TYPE: prt
16 ORGANISM: Homo sapiens
17 US-09-955-807-5

```

Query Match	28.3%	Score 51	DB 10	Length 65
Best Local Similarity	52.2%	Pred. No. 17		
Matches 12	Conservative 3	Mismatches 4	Indels 4	Gaps 1

```

QY      9 PSTGCFSG---KETPSDDRQ 27
          | : ||| | : |||
Db     14 PAVTGTGDSGAKETVSDKRSQ 36

```

```

RESULT 6
US-09-955-807-10
: Sequence 10 Application US/09955807
: Patent No. US20020132996A1
: GENERAL INFORMATION:
: APPLICANT: Lok, Si
: APPLICANT: Shepard, Paul O.
: APPLICANT: Kinsdavegl, Wayne
: APPLICANT: Bort, Susan J.
: TITLE OF INVENTION: Secretary Protein-48
: FILE REFERENCE: 98-17C1
: CURRENT APPLICATION NUMBER: US/09/955,807
: PRIOR FILING DATE: 2001-09-19
: PRIOR APPLICATION NUMBER: 60/102,679
: PRIOR FILING DATE: 1998-10-01
: PRIOR APPLICATION NUMBER: 09/410,603
: PRIOR FILING DATE: 1999-10-01
: NUMBER OF SEQ ID NOS: 17
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 10
: LENGTH: 65
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-955-807-10

```

Query Match	28.3%	Score 51;	DB 10;	Length 65;
Best Local Similarity	52.2%	Pred. No. 17;		
Matches 12;	Conservative 3;	Mismatches 4;	Indels 4;	Gaps 1;

QY 9 PSTGCGSG---KETPSDDRSQ 27
| : ||| ||| ||| |||
Db 17 PAVTGFSGDSGAKETVSQDKRSQ 39

RESULT 7
US-09-955-807-4

; Sequence 4, Application US/09955807
; Patent No. US20020132996A1

GENERAL INFORMATION:

APPLICANT: Sheppard, Paul O.

APPLICANT: Kindsvogel, Wayne

TITLE OF INVENTION: Secretory Protein-48

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

```

?
? CURRENT FILING DATE: 2001-09-19
?
? PRIOR APPLICATION NUMBER: 60/102,679
?
? PRIOR FILING DATE: 1998-10-01
?
? PRIOR APPLICATION NUMBER: 09/410,603
?
? PRIOR FILING DATE: 1999-10-01
?
? NUMBER OF SEQ ID NOS: 17
?
? SOFTWARE: FastSeq for Windows Version 3.0
?
? SEQ ID NO 4
?
? LENGTH: 77
?
? TYPE: prt
?
? ORGANISM: Homo sapiens
?
? US-09-955-807-4

```

Query Match	28.3%	Score 51;	DB 10;	Length 77;
Best Local Similarity	52.2%	Pred. NO. 20;		
Matches 12; Conservative	3;	Mismatches 4;	Indels 4;	Gaps 1.

```

QY      9 PSTGFGSG---KETPSDDRSG 27
          | : | | | | | |
Db     26 PAVTGTGDSGAKETVSDKRSQ 48

```

```

RESULT 8
US-09-955-807-3
: Sequence 3, Application US/099555807
: Patent No. US2002012996A1
: GENERAL INFORMATION:
: APPLICANT: Lok, Si
: APPLICANT: Sheppard, Paul O.
: APPLICANT: Kindsvogel, Wayne
: APPLICANT: Bort, Susan J
: TITLE OR INVENTION: Secretary Protein-48
: FILE REFERENCE: 98-17C1
: CURRENT APPLICATION NUMBER: US/09/955,807
: PRIORITY FILING DATE: 2001-09-19
: PRIOR APPLICATION NUMBER: 60/102,679
: PRIOR FILING DATE: 1998-10-01
: PRIOR APPLICATION NUMBER: 09/410,603
: PRIOR FILING DATE: 1999-10-01
: NUMBER OF SEQ ID NOS: 17
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 3
: LENGTH: 79
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-955-807-3

```

Query Match	28.3%	Score 51;	DB 10;	Length 79;
Best Local Similarity	52.2%	Pred. No. 21;		
Matches 12;	Conservative 3;	Mismatches 4;	Indels 4;	Gaps 1;

```

Qy      9 PSTGFGSG---KETPSDDRSG 27
          | : | | | | | |
Db      28 PAVTGTGDSGAKETVSDKRSQ 50

```

RESULT 9
US-09-955-807-16

Patent No. US2002013

GENERAL INFORMATION:

APPLICANT: Sheppard, Paul O.

APPLICANT: Bort, Susan J.

FILE REFERENCE: 98-17C1

CURRENT APPLICATION NUMBER: US/09/955,807

PRIOR APPLICATION NUMBER: 60/102,679

PRIOR APPLICATION NUMBER: 09/410,603

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

```
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-955-807-16

Query Match          28.3%; Score 51; DB 10; Length 80;
Best Local Similarity 52.2%; Pred. No. 21;
Matches 12; Conservative 3; Mismatches 4; Indels 4; Gaps 1;

QY      9 PSTGFGSG---KETPSEDDRSQ 27
Db      29 PAVTGFTGDSGAKETVSQDKRSQ 51

RESULT 10
US-09-955-807-15
; Sequence 15, Application US/09955807
; Patent No. US2002013296A1
; GENERAL INFORMATION:
; APPLICANT: Lok, Si
; APPLICANT: Shepard, Paul O.
; APPLICANT: Kindsvogel, Wayne
; APPLICANT: Bort, Susan J.
; TITLE OF INVENTION: Secretory Protein-48
; FILE REFERENCE: 98-17C1
; CURRENT APPLICATION NUMBER: US/09/955,807
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/102,679
; PRIOR FILING DATE: 1998-10-01
; PRIOR APPLICATION NUMBER: 09/410,603
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 89
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-955-807-15

Query Match          28.3%; Score 51; DB 10; Length 89;
Best Local Similarity 52.2%; Pred. No. 24;
Matches 12; Conservative 3; Mismatches 4; Indels 4; Gaps 1;

QY      9 PSTGFGSG---KETPSEDDRSQ 27
Db      38 PAVTGFTGDSGAKETVSQDKRSQ 60

RESULT 11
US-09-955-807-2
; Sequence 2, Application US/09955807
; Patent No. US2002013296A1
; GENERAL INFORMATION:
; APPLICANT: Lok, Si
; APPLICANT: Shepard, Paul O.
; APPLICANT: Kindsvogel, Wayne
; APPLICANT: Bort, Susan J.
; TITLE OF INVENTION: Secretory Protein-48
; FILE REFERENCE: 98-17C1
; CURRENT APPLICATION NUMBER: US/09/955,807
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/102,679
; PRIOR FILING DATE: 1998-10-01
; PRIOR APPLICATION NUMBER: 09/410,603
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 105
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
US-09-955-807-2

Query Match          28.3%; Score 51; DB 10; Length 105;
Best Local Similarity 52.2%; Pred. No. 28;
Matches 12; Conservative 3; Mismatches 4; Indels 4; Gaps 1;

QY      9 PSTGFGSG---KETPSEDDRSQ 27
Db      54 PAVTGFTGDSGAKETVSQDKRSQ 76

RESULT 12
US-09-955-807-14
; Sequence 14, Application US/09955807
; Patent No. US2002013296A1
; GENERAL INFORMATION:
; APPLICANT: Lok, Si
; APPLICANT: Shepard, Paul O.
; APPLICANT: Kindsvogel, Wayne
; APPLICANT: Bort, Susan J.
; TITLE OF INVENTION: Secretory Protein-48
; FILE REFERENCE: 98-17C1
; CURRENT APPLICATION NUMBER: US/09/955,807
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/102,679
; PRIOR FILING DATE: 1998-10-01
; PRIOR APPLICATION NUMBER: 09/410,603
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 105
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-955-807-14

Query Match          28.3%; Score 51; DB 10; Length 105;
Best Local Similarity 52.2%; Pred. No. 28;
Matches 12; Conservative 3; Mismatches 4; Indels 4; Gaps 1;

QY      9 PSTGFGSG---KETPSEDDRSQ 27
Db      54 PAVTGFTGDSGAKETVSQDKRSQ 76

RESULT 13
US-09-904-987-5
; Sequence 5, Application US/09904987
; Patent No. US20020037908A1
; GENERAL INFORMATION:
; APPLICANT: No. US20020037908A1actyl, Inc.
; TITLE OF INVENTION: Methods and Compositions for Controlling Pathological and Prepath
; FILE REFERENCE: 42108/26146
; CURRENT APPLICATION NUMBER: US/09/904,987
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 758
; TYPE: PRT
; ORGANISM: homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NCBI ENTREZ / NM_016835
; DATABASE ENTRY DATE: 2001-02-13
; RELEVANT RESIDUES: (1)..(758)
US-09-904-987-5

Query Match          28.3%; Score 51; DB 9; Length 758;
Best Local Similarity 27.3%; Pred. No. 2.5e+02;
Matches 12; Conservative 7; Mismatches 13; Indels 12; Gaps 1;
```

```

Oy      4 SESSTPSTGSPGSKETPSE-----DNRSGSHNGGES 35
Db      282 SEPQPSVGRAKGQDAPLEFFHVEITPNVQKQKQSHSEHIGRA 325

RESULT 14
US-09-764-891-4686
; Sequence 4686, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4686
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (62)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (197)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (205)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-891-4686

Query Match      27.8%; Score 50; DB 11; Length 212;
Best Local Similarity 35.7%; Pred. No. 82;
Matches 10; Conservative 4; Mismatches 14; Indels 0; Gaps 0;

Oy      1 SFOSESSTPSTGSPGSKETPSEDDRSQS 28
Db      161 NLSNWLFSPQCGGLTGKAKKEDGTSSR 188

RESULT 15
US-09-864-761-37733
; Sequence 37733, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wenheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30

```

```

; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 37733
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006059.3
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN PETAL LIVER, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.8
; OTHER INFORMATION: EST HUMAN HIT: A1350947.1, EVALUATE 1.00e-78
; OTHER INFORMATION: SWISSPROT HIT: P30414, EVALUATE 1.00e-108
US-09-864-761-37733

Query Match      27.8%; Score 50; DB 9; Length 253;
Best Local Similarity 66.7%; Pred. No. 99;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Oy      15 SSKETPSEDDRSQSR 29
Db      203 SDEQTPSRDDDSQSR 217

Search completed: July 24, 2003, 12:18:18
Job time : 6.24832 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: July 24, 2003, 11:49:53 ; Search time 52.7306 Seconds
(without alignments)
577.701 Million cell updates/sec

Title: US-09-991-681-29

Perfect score: 180
Sequence: 1 SFQSSSTPTSGFGSKRPSEDDRSQSRHMGES 35

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 5580241 seqs, 870357830 residues

Total number of hits satisfying chosen parameters: 5580241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA_Main:*

- 1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
- 2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
- 3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
- 4: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
- 5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
- 6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
- 7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
- 8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
- 9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
- 10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
- 11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
- 12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
- 13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
- 14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
- 15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
- 16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
- 17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
- 18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
- 19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
- 20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
- 21: /cgn2_6/ptodata/1/paa/US097A_COMB.pep.*
- 22: /cgn2_6/ptodata/1/paa/US097B_COMB.pep.*
- 23: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
- 24: /cgn2_6/ptodata/1/paa/US099A_COMB.pep.*
- 25: /cgn2_6/ptodata/1/paa/US099B_COMB.pep.*
- 26: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
- 27: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
- 28: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
- 29: /cgn2_6/ptodata/1/paa/US103_COMB.pep.*
- 30: /cgn2_6/ptodata/1/paa/US104_COMB.pep.*
- 31: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	180	100.0	35	12 US-08-842-385-8	Sequence 8, Appl1
2	180	100.0	35	25 US-09-991-681-29	Sequence 29, Appl1

3	180	100.0	467	12 US-08-842-385-6	Sequence 6, Appl1
4	180	100.0	518	25 US-09-991-681-27	Sequence 27, Appl1
5	180	100.0	1770	1 PCT-US03-01943-44	Sequence 44, Appl1
6	180	100.0	1770	US-10-144-198-44	Sequence 44, Appl1
7	180	100.0	1807	1 PCT-US01-08631-40087	Sequence 4087, A
8	180	100.0	1839	1 PCT-US01-42950-495	Sequence 495, App
9	180	100.0	1839	US-10-416-993-495	Sequence 495, App
10	180	100.0	1872	1 PCT-US03-04508-32	Sequence 32, Appl1
11	180	100.0	1982	1 PCT-US01-08631-40090	Sequence 4090, A
12	180	100.0	2221	1 PCT-US03-01943-30	Sequence 30, Appl1
13	180	100.0	2221	US-10-144-198-30	Sequence 30, Appl1
14	143	79.4	141	28 US-10-221-279-7783	Sequence 7783, Ap
15	61	33.9	1061	20 US-09-614-150-4746	Sequence 4746, Ap
16	61	33.9	1061	31 US-60-167-217-4862	Sequence 4862, Ap
17	61	33.9	1061	31 US-60-173-464-3932	Sequence 3932, Ap
18	61	33.9	1061	31 US-60-191-637-4762	Sequence 4762, Ap
19	61	33.9	1061	31 US-60-191-681-3762	Sequence 3762, Ap
20	59	32.8	253	1 PCT-US01-08631-37316	Sequence 37316, A
21	59	32.8	1378	31 US-60-167-217-8500	Sequence 8500, Ap
22	59	32.8	1403	20 US-09-614-150-8412	Sequence 8412, Ap
23	59	32.8	1403	31 US-60-173-464-6811	Sequence 6811, Ap
24	59	32.8	1403	31 US-60-191-637-8436	Sequence 8436, Ap
25	59	32.8	1403	31 US-60-191-681-6532	Sequence 6532, Ap
26	58	32.2	73	1 PCT-US02-32727-4001	Sequence 4001, Ap
27	58	32.2	73	25 US-09-978-825-4001	Sequence 4001, Ap
28	58	32.2	73	26 US-10-057-498-4001	Sequence 4001, Ap
29	58	32.2	245	28 US-10-219-999-55380	Sequence 55380, A
30	58	32.2	245	30 US-10-425-114-66127	Sequence 66127, A
31	58	32.2	366	21 US-09-708-427-73150	Sequence 73150, A
32	58	32.2	514	30 US-10-437-963-119781	Sequence 119781, A
33	57.5	31.9	592	22 US-09-791-537-48904	Sequence 48904, A
34	57.5	31.9	765	30 US-10-438-246-10453	Sequence 10453, A
35	57	31.7	230	30 US-10-438-246-30307	Sequence 30307, A
36	57	31.7	294	31 US-60-312-544-6563	Sequence 6563, Ap
37	57	31.7	345	20 US-09-614-150-37497	Sequence 37497, A
38	57	31.7	345	31 US-60-173-464-28442	Sequence 28442, A
39	57	31.7	345	31 US-60-191-637-37111	Sequence 37111, A
40	57	31.7	345	31 US-60-191-681-28918	Sequence 28918, A
41	57	31.7	388	27 US-10-155-881-18359	Sequence 18359, A
42	57	31.7	410	27 US-10-155-881-31432	Sequence 31432, A
43	57	31.7	410	30 US-10-425-114-40793	Sequence 40793, A
44	57	31.7	410	31 US-60-312-544-7974	Sequence 7974, Ap
45	57	31.7	528	27 US-10-155-881-18358	Sequence 18358, A

ALIGNMENTS

RESULT 1
US-08-842-385-8
Sequence 8, Application US/08842385

GENERAL INFORMATION:
APPLICANT: Russell, John
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL FOR DETECTING DISEASE OF THE PROSTATE
TITLE OF INVENTION: FOR DETECTING DISEASE OF THE PROSTATE
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESSES:
ADDRESS: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/842,385
FILING DATE:
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Foremski, Priscilla E
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 6084.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/937-6365
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 35 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: None
US-08-842-385-8

Query Match 100.0%; Score 180; DB 12;
Best Local Similarity 100.0%; Pred. No. 9.3e-16;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SFQSSSTPSTGFGSGKETPSDDRSQSRHMGES 35
Db 1 SFQSSSTPSTGFGSGKETPSDDRSQSRHMGES 35

RESULT 2
US-09-991-681-29
Sequence 29, Application US/09991681
GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA
COHEN, MAURICE
COLPITTS, TRACEY L.
FRIEDMAN, PAULA N.
GORDON, JULIAN
GRANADOS, EDWARD N.
HODGES, STEVEN C.
KLASS, MICHAEL R.
KRATOCHVIL, JON D.
ROBERTS-RAPP, LISA
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSER: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/991,681
FILING DATE: 26-Nov-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/065,383
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6084.US.PI
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 35 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: None
SEQUENCE DESCRIPTION: SEQ ID NO: 29:
US-09-991-681-29

Query Match 100.0%; Score 180; DB 25; Length 467;
Best Local Similarity 100.0%; Pred. No. 2.4e-14;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SFQSSSTPSTGFGSGKETPSDDRSQSRHMGES 35
Db 1 SFQSSSTPSTGFGSGKETPSDDRSQSRHMGES 35

RESULT 3
US-08-842-385-6
Sequence 6, Application US/08842385
GENERAL INFORMATION:
APPLICANT: Russell, John
COLPITTS, TRACEY
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
FOR DETECTING DISEASE OF THE PROSTATE
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSER: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/842,385
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Foremski, Priscilla E
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 6084.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/937-6365
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: None
US-08-842-385-6

Query Match 100.0%; Score 180; DB 12; Length 467;
Best Local Similarity 100.0%; Pred. No. 2.4e-14;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SFQSSSTPSTGFGSGKETPSDDRSQSRHMGES 35
Db 232 SFQSSSTPSTGFGSGKETPSDDRSQSRHMGES 266

```
RESULT 4
US-09-991-681-27
; Sequence 27, Application US/09991681
; GENERAL INFORMATION:
; APPLICANT: BILLING-MEDEL, PATRICIA
; COHEN, MAURICE
; COLPITTS, TRACEY L.
; FRIEDMAN, PAULA N.
; GORDON, JULIAN
; GRANADOS, EDWARD N.
; HODGES, STEVEN C.
; KLAAS, MICHAEL R.
; KRATOCHVIL, JON D.
; ROBERTS-RAPP, LISA
; TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
; FOR DETECTING DISEASES OF THE PROSTATE
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/991,681
; FILING DATE: 26-Nov-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/065,383
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Becker, Cheryl L.
; REGISTRATION NUMBER: 35,441
; REFERENCE/DOCKET NUMBER: 6084.US.P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847/935-1729
; TELEFAX: 847/935-2623
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 518 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: None
; SEQUENCE DESCRIPTION: SEQ ID NO: 27:
US-09-991-681-27

Query Match 100.0%; Score 180; DB 25; Length 518;
Best Local Similarity 100.0%; Pred. No. 2.8e-14;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 SFOSSSTPSTGSGKTPSEDDRSQSRHMGES 35
Db 283 SFOSSSTPSTGSGKTPSEDDRSQSRHMGES 317

RESULT 5
PCT-US03-01943-44
; Sequence 44, Application PC/TUS0301943
; GENERAL INFORMATION:
; APPLICANT: ORIGENE TECHNOLOGIES INC
; TITLE OF INVENTION: CANCER GENES
; FILE REFERENCE: 3U 9U 901 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/01943
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: US 10/054,935
```

```
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: US 60/356,130
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 10/102,946
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: US 10/117,229
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: US 10/144,198
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: US 10/197,824
; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44
; LENGTH: 1770
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-01943-44

Query Match 100.0%; Score 180; DB 1; Length 1770;
Best Local Similarity 100.0%; Pred. No. 1.3e-13;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 SFOSSSTPSTGSGKTPSEDDRSQSRHMGES 35
Db 1535 SFOSSSTPSTGSGKTPSEDDRSQSRHMGES 1569

RESULT 6
US-10-144-198-44
; Sequence 44, Application US/10144198
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies Inc
; TITLE OF INVENTION: Regulated Prostate Cance Genes
; FILE REFERENCE: 9U 105 R1
; CURRENT APPLICATION NUMBER: US/10/144,198
; CURRENT FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 44
; LENGTH: 1770
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-198-44

Query Match 100.0%; Score 180; DB 27; Length 1770;
Best Local Similarity 100.0%; Pred. No. 1.3e-13;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 SFOSSSTPSTGSGKTPSEDDRSQSRHMGES 35
Db 1535 SFOSSSTPSTGSGKTPSEDDRSQSRHMGES 1569

RESULT 7
PCT-US01-08631-40087
; Sequence 40087, Application PC/TUS0108631
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-049
; CURRENT APPLICATION NUMBER: PCT/US01/08631
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 40087
; LENGTH: 1807
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (48)..(62)
OTHER INFORMATION: Fusion glycoprotein F0 domain identified by eMATRIX,
OTHER INFORMATION: accession number PF00523D, p-value=7.188e-10, raw score of 11.39
NAME/KEY: DOMAIN
LOCATION: (941)..(950)
OTHER INFORMATION: Helper component proteinase domain identified by Pfam,
OTHER INFORMATION: accession name Peptidase_C6, E-value=0.0056, Pfam score of 7.9
PCT-US01-08631-40087
```

```
Query Match          100.0%; Score 180; DB 1; Length 1807;
Best Local Similarity 100.0%; Pred. No. 1.3e-13;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 SFQSSSTPSTGFGSGKETPSEDDRSQSREHMGES 35
Db 1572 SFQSSSTPSTGFGSGKETPSEDDRSQSREHMGES 1606
```

```
RESULT 8
PCT-US01-42950-495
Sequence 495, Application PC/TUS0142950
GENERAL INFORMATION:
```

```
APPLICANT: HySeq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 21272-096
CURRENT APPLICATION NUMBER: PCT/US01/42950
PRIOR FILING DATE: 2001-11-16
PRIOR APPLICATION NUMBER: 09/714,936
PRIOR FILING DATE: 2000-11-17
NUMBER OF SEQ ID NOS: 682
SOFTWARE: PatentIn version 3.0
SEQ ID NO 495
LENGTH: 1839
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US01-42950-495
```

```
Query Match          100.0%; Score 180; DB 1; Length 1839;
Best Local Similarity 100.0%; Pred. No. 1.4e-13;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 SFQSSSTPSTGFGSGKETPSEDDRSQSREHMGES 35
Db 1604 SFQSSSTPSTGFGSGKETPSEDDRSQSREHMGES 1638
```

```
RESULT 9
US-10-416-993-495
Sequence 495, Application US/10416993
GENERAL INFORMATION:
```

```
APPLICANT: HySeq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 21272-096
CURRENT APPLICATION NUMBER: US/10/416,993
PRIOR FILING DATE: 2003-11-16
PRIOR APPLICATION NUMBER: 09/714,936
PRIOR FILING DATE: 2000-11-17
NUMBER OF SEQ ID NOS: 682
SOFTWARE: PatentIn version 3.0
SEQ ID NO 495
LENGTH: 1839
TYPE: PRT
ORGANISM: Homo sapiens
US-10-416-993-495
```

```
Query Match          100.0%; Score 180; DB 30; Length 1839;
Best Local Similarity 100.0%; Pred. No. 1.4e-13;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 SFQSSSTPSTGFGSGKETPSEDDRSQSREHMGES 35
Db 1572 SFQSSSTPSTGFGSGKETPSEDDRSQSREHMGES 1606
```

```
Db 1604 SFQSSSTPSTGFGSGKETPSEDDRSQSREHMGES 1638
```

```
RESULT 10
PCT-US03-04508-32
Sequence 32, Application PC/TUS0304508
GENERAL INFORMATION:
```

```
APPLICANT: IDEC PHARMACEUTICALS
APPLICANT: GATELY, DENNIS
TITLE OF INVENTION: PROSTATE SPECIFIC GENES AND THE USE THEREOF IN DESIGN
FILE REFERENCE: 037003/0301985
CURRENT APPLICATION NUMBER: PCT/US03/04508
CURRENT FILING DATE: 2003-02-19
PRIOR APPLICATION NUMBER: 60/357,140
PRIOR FILING DATE: 2002-02-19
PRIOR APPLICATION NUMBER: 60/396,082
PRIOR FILING DATE: 2002-07-17
PRIOR APPLICATION NUMBER: 60/386,759
PRIOR FILING DATE: 2002-06-10
NUMBER OF SEQ ID NOS: 93
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 32
LENGTH: 1872
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US03-04508-32
```

```
Query Match          100.0%; Score 180; DB 1; Length 1872;
Best Local Similarity 100.0%; Pred. No. 1.4e-13;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 SFQSSSTPSTGFGSGKETPSEDDRSQSREHMGES 35
Db 1637 SFQSSSTPSTGFGSGKETPSEDDRSQSREHMGES 1671
```

```
RESULT 11
PCT-US01-08631-40090
Sequence 40090, Application PC/TUS0108631
GENERAL INFORMATION:
```

```
APPLICANT: HySeq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 21272-049
CURRENT APPLICATION NUMBER: PCT/US01/08631
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 09/540,217
PRIOR FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: 09/649,167
PRIOR FILING DATE: 2000-08-23
NUMBER OF SEQ ID NOS: 60736
SOFTWARE: Custom
SEQ ID NO 40090
LENGTH: 1982
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (11)..(25)
OTHER INFORMATION: Fusion glycoprotein F0 domain identified by eMATRIX,
OTHER INFORMATION: accession number PF00523D, p-value=7.188e-10, raw score of 11.39
NAME/KEY: DOMAIN
LOCATION: (1065)..(1074)
OTHER INFORMATION: Helper component proteinase domain identified by Pfam,
OTHER INFORMATION: accession name Peptidase_C6, E-value=0.0056, Pfam score of 7.9
PCT-US01-08631-40090
```

```
Query Match          100.0%; Score 180; DB 1; Length 1982;
Best Local Similarity 100.0%; Pred. No. 1.5e-13;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 SFQSSSTPSTGFGSGKETPSEDDRSQSREHMGES 35
Db 1572 SFQSSSTPSTGFGSGKETPSEDDRSQSREHMGES 1606
```



```
Db      1686 SFOSESTPSTGCGKTPSEDDRSQSRHMGES 1720
RESULT 12
PCT-US03-01943-30
; Sequence 30, Application PC/TUS0301943
; GENERAL INFORMATION:
; APPLICANT: ORIGENE TECHNOLOGIES INC
; TITLE OF INVENTION: CANCER GENES
; FILE REFERENCE: 3U 9U 901 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/01943
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: US 10/054,935
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: US 60/356,130
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 10/102,946
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: US 10/117,229
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: US 10/144,198
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: US 10/197,824
; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30
; LENGTH: 2221
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-01943-30

Query Match      100.0%; Score 180; DB 1; Length 2221;
Best Local Similarity 100.0%; Pred. No. 1.7e-13;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 SFOSESTPSTGCGKTPSEDDRSQSRHMGES 35
Db      1986 SFOSESTPSTGCGKTPSEDDRSQSRHMGES 2020
RESULT 13
US-10-144-198-30
; Sequence 30, Application US/10144198
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies Inc
; TITLE OF INVENTION: Regulated Prostate Cance Genes
; FILE REFERENCE: 9U 105 R1
; CURRENT APPLICATION NUMBER: US/10/144,198
; CURRENT FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 30
; LENGTH: 2221
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-198-30

Query Match      100.0%; Score 180; DB 27; Length 2221;
Best Local Similarity 100.0%; Pred. No. 1.7e-13;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 SFOSESTPSTGCGKTPSEDDRSQSRHMGES 35
Db      1986 SFOSESTPSTGCGKTPSEDDRSQSRHMGES 2020

RESULT 14
US-10-221-279-7783
; Sequence 7783, Application US/10221279
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides

Db      1686 SFOSESTPSTGCGKTPSEDDRSQSRHMGES 1720
RESULT 15
US-09-614-150-4746
; Sequence 4746, Application US/09614150
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CLO00728
; CURRENT APPLICATION NUMBER: US/09/614,150
; CURRENT FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4746
; LENGTH: 1061
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-09-614-150-4746

Query Match      33.9%; Score 61; DB 20; Length 1061;
Best Local Similarity 51.9%; Pred. No. 2.3e+02;
Matches 14; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Qy      4 SSSSTPSTGCGKTPSEDDRSQSRHMGES 30
Db      39 SKSRPSSAGVVIDETQSEEBESQSSSE 65
```

Fri Jul 25 15:44:27 2003

us-09-991-681-29.rapm

Page 6

Search completed: July 24, 2003, 12:15:13
Job time : 53.7306 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 11:54:33 ; Search time 1.02489 Seconds
(without alignments)
140.227 Million cell updates/sec

Title: US-09-991-681-29

Perfect score: 180
Sequence: 1 SFOSESSTPTGFGFGKTPEDDRSOSREHMGES 35

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 41799 seqs, 4106219 residues

Total number of hits satisfying chosen parameters: 41799

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA New:*
1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	100	100.0	1872	US-10-367-978-32	Sequence 32, Appl
2	56	31.1	2047	US-10-370-480-28	Sequence 28, Appl
3	56	31.1	5596	US-60-479-073-296	Sequence 296, Appl
4	50.5	28.1	233	US-10-273-573-9796	Sequence 9796, Appl
5	50	27.8	1419	PCT-US02-35624-173	Sequence 173, Appl
6	49	27.2	817	US-10-273-573-8806	Sequence 8806, Appl
7	48	26.7	402	US-10-294-433-286	Sequence 286, Appl
8	47	26.1	301	US-10-273-573-10678	Sequence 10678, A
9	47	26.1	1189	US-10-273-573-8991	Sequence 8991, Ap
10	46.5	25.8	743	US-60-478-196-3144	Sequence 3144, Ap
11	46.5	25.8	186	PCT-US02-18638A-188	Sequence 188, Appl
12	46	25.6	186	US-10-372-209-12	Sequence 12, Appl
13	46	25.6	282	US-10-372-209-13	Sequence 13, Appl
14	46	25.6	282	US-10-372-209-18	Sequence 18, Appl
15	46	25.6	565	US-10-380-565-34	Sequence 34, Appl
16	45	25.0	137	US-10-273-573-6227	Sequence 6227, Ap
17	45	25.0	137	US-10-273-573-10834	Sequence 10834, A
18	45	25.0	442	US-10-294-433-293	Sequence 293, Appl
19	44.5	24.7	174	US-10-273-573-6423	Sequence 6423, Ap
20	44.5	24.7	617	US-10-294-433-396	Sequence 396, Appl
21	44.5	24.7	979	PCT-US03-19743-2	Sequence 2, Appl
22	44	24.4	47	US-10-323-069A-135	Sequence 135, Appl
23	44	24.4	306	US-10-294-433-841	Sequence 841, Appl
24	44	24.4	450	US-10-273-573-8451	Sequence 8451, Appl
25	44	24.4	915	US-10-331-496A-68	Sequence 68, Appl
26	44	24.4	967	US-10-331-496A-95	Sequence 95, Appl

27	43.5	24.2	250	US-10-273-573-7905	Sequence 7905, Ap
28	43.5	24.2	846	US-60-478-196-3082	Sequence 3082, Ap
29	43	23.9	937	US-60-478-196-3207	Sequence 3207, Ap
30	43	23.9	1994	US-10-294-433-339	Sequence 339, Appl
31	43	23.9	2041	US-10-294-433-338	Sequence 338, Appl
32	42.5	23.6	131	US-10-273-573-7534	Sequence 7534, Appl
33	42.5	23.6	131	US-10-273-573-7876	Sequence 7876, Appl
34	42.5	23.6	402	US-09-635-359B-19	Sequence 19, Appl
35	42.5	23.6	439	US-60-478-196-3261	Sequence 3261, Appl
36	42.5	23.6	609	US-60-479-073-439	Sequence 439, Appl
37	42.5	23.6	753	US-60-479-073-437	Sequence 437, Appl
38	42.5	23.6	793	PCT-US03-11867-4	Sequence 4, Appl
39	42	23.3	74	US-10-273-573-7515	Sequence 7515, Appl
40	42	23.3	140	US-10-273-573-10371	Sequence 10371, A
41	42	23.3	630	US-60-478-196-3344	Sequence 3344, Appl
42	42	23.3	787	US-10-372-327-43	Sequence 43, Appl
43	42	23.3	1627	US-10-433-757-5	Sequence 5, Appl
44	41.5	23.1	152	US-10-294-433-849	Sequence 849, Appl
45	41.5	23.1	282	US-10-273-573-8747	Sequence 8747, Appl

ALIGNMENTS

```

RESULT 1
US-10-367-978-32
Sequence 32, Application US/10367978
GENERAL INFORMATION:
APPLICANT: GATELY, DENNIS
TITLE OF INVENTION: PROSTATE SPECIFIC GENES AND THE USE THEREOF IN DESIGN
FILE REFERENCE: 037003-0301988
CURRENT APPLICATION NUMBER: US/10/367,978
PRIOR APPLICATION NUMBER: 60/357,140
PRIOR FILING DATE: 2002-02-19
PRIOR APPLICATION NUMBER: 60/396,082
PRIOR FILING DATE: 2002-07-17
PRIOR APPLICATION NUMBER: 60/386,759
PRIOR FILING DATE: 2002-06-10
NUMBER OF SEQ ID NOS: 89
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 32
LENGTH: 1872
TYPE: PRT
ORGANISM: Homo sapiens
US-10-367-978-32

Query Match          100.0%; Score 180; DB 6; Length 1872;
Best Local Similarity 100.0%; Pred. No. 1.4e-16;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1 SFOSESSTPTGFGFGKTPEDDRSOSREHMGES 35
Db      1637 SFOSESSTPTGFGFGKTPEDDRSOSREHMGES 1671

RESULT 2
US-10-370-480-28
Sequence 28, Application US/10370480
GENERAL INFORMATION:
APPLICANT: Hitachi Ltd.
TITLE OF INVENTION: ADP INTERACTING PROTEINS AND USE THEREOF
FILE REFERENCE: HITA.0167
CURRENT APPLICATION NUMBER: US/10/370,480
PRIOR FILING DATE: 2003-02-24
PRIOR APPLICATION NUMBER: US 60/358,319
PRIOR FILING DATE: 2002-02-22
PRIOR APPLICATION NUMBER: US 60/407,252
NUMBER OF SEQ ID NOS: 45
SOFTWARE: PatentIn version 3.1
SEQ ID NO 28

```

```

; LENGTH: 2047
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-370-480-28

Query Match
Best Local Similarity 31.1%; Score 56; DB 6; Length 2047;
Matches 12; Conservative 6; Mismatches 8; Indels 4; Gaps 1;

Qy 3 QSESSTPSTGFGSGKETPSEDSDRSQSRHMG 32
Db 1383 QNESOSQP-----EPREGPSEDKABGEHEM 1408

RESULT 3
US-60-479-073-296
; Sequence 296, Application US/60479073
; GENERAL INFORMATION:
; APPLICANT: De Wilde, Gert Jules Hector
; APPLICANT: Saudeberg, Michael John Scott
; APPLICANT: Logghe, Marc Georges
; TITLE OF INVENTION: Amino acid sequences useful for developing compounds for the
; TITLE OF INVENTION: prevention and/or treatment of metabolic diseases and nucleotide
; FILE REFERENCE: D00590.70042.US
; CURRENT APPLICATION NUMBER: US/60/479,073
; CURRENT FILING DATE: 2003-06-17
; NUMBER OF SEQ ID NOS: 526
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 296
; LENGTH: 5596
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-479-073-296

Query Match
Best Local Similarity 31.1%; Score 56; DB 7; Length 5596;
Matches 12; Conservative 6; Mismatches 8; Indels 4; Gaps 1;

Qy 3 QSESSTPSTGFGSGKETPSEDSDRSQSRHMG 32
Db 4932 QNESOSQP-----EPREGPSEDKABGEHEM 4957

RESULT 4
US-10-273-573-9796
; Sequence 9796, Application US/10273573
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: US/10/273,573
; CURRENT FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 9796
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-273-573-9796

Query Match
Best Local Similarity 28.1%; Score 50.5; DB 6; Length 233;
Matches 13; Conservative 3; Mismatches 14; Indels 9; Gaps 1;

Qy 5 ESSTPSTGFG-----SGKETPSEDSDRSQSRHMG 34
Db 33 EPPTPDGGGCVPRSGSRPRDAQEAPVPDIRSQHKVHTNE 71
```

```

RESULT 5
PCT-US02-35624-173
; Sequence 173, Application PC/TUS0235624
; GENERAL INFORMATION:
; APPLICANT: Corixa Corporation
; APPLICANT: Maisonneuve, Jean-Francois L.
; APPLICANT: Barth, Brenda
; APPLICANT: Bhatia, Ajay
; APPLICANT: Probst, Peter
; APPLICANT: Gunderian, Jeff
; APPLICANT: Skeiky, Yasir A. W.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
; TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.51501PC
; CURRENT APPLICATION NUMBER: PCT/US02/35624
; CURRENT FILING DATE: 2003-05-30
; NUMBER OF SEQ ID NOS: 175
; SEQ ID NO 173
; LENGTH: 1419
; TYPE: PRT
; ORGANISM: Chlamydia
PCT-US02-35624-173

Query Match
Best Local Similarity 27.8%; Score 50; DB 1; Length 1419;
Matches 12; Conservative 5; Mismatches 18; Indels 0; Gaps 0;

Qy 1 SFOESSSTPSTGFGSGKETPSEDSDRSQSRHMG 35
Db 704 SSESSTPSTGFGSGKETPSEDSDRSQSRHMG 738

RESULT 6
US-10-273-573-8906
; Sequence 8906, Application US/10273573
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: US/10/273,573
; CURRENT FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 8906
; LENGTH: 817
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: DOMAIN
; LOCATION: (478)..(498)
; OTHER INFORMATION: Elongation factor 1 beta/beta'/delta chain proteins domain
; OTHER INFORMATION: Identified by eMATRIX, accession number BL00624B, p-value=7.130e-
; OTHER INFORMATION: 10, raw score of 9.21
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(817)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-273-573-8906

Query Match
Best Local Similarity 27.2%; Score 49; DB 6; Length 817;
Matches 16; Conservative 1; Mismatches 13; Indels 10; Gaps 2;

Qy 4 SESSTPSTGFGSGKET-----PSDDDRSQSRHMG 34
Db 329 SESSEPGKGPASDSSTEDASRHKQPSDDDS-DRENKGE 367
```

```
RESULT 7
US-10-294-433-286
; Sequence 286, Application US/10294433
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 792C1P4
; CURRENT APPLICATION NUMBER: US/10/294,433
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: PCT/US01/14826
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/989,600
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 09/577,408
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: 10/115,831
; PRIOR FILING DATE: 2002-04-02
; PRIOR APPLICATION NUMBER: 09/677,298
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/695,781
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 10/150,802
; PRIOR FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: 09/715,869
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 10/167,379
; PRIOR FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 09/775,330
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 864
; SOFTWARE: Custom
; SEQ ID NO 286
; LENGTH: 402
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-294-433-286

Query Match      26.1%; Score 48; DB 6; Length 402;
Best Local Similarity 40.0%; Pred. No. 13;
Matches 12; Conservative 5; Mismatches 11; Indels 2; Gaps 1;

QY      4 SESSTPSTGCGSGKET--PSRDDRSQSRHH 31
DB      314 SRAQFPSTGCGSGYKNNPGEMRRARRKRKH 343

RESULT 8
US-10-273-573-10678
; Sequence 10678, Application US/10273573
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: US/10/273,573
; PRIOR FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 10678
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-273-573-10678

Query Match      26.1%; Score 47; DB 6; Length 301;
Best Local Similarity 38.5%; Pred. No. 13;
Matches 10; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY      4 SESSTPSTGCGSGKETPSRDDRSQSR 29
DB      314 SRAQFPSTGCGSGYKNNPGEMRRARRKRKH 343
```

```
DB      77 ARSNAPTAGGRRRRRREDVRSRAR 102

RESULT 9
US-10-273-573-8991
; Sequence 8991, Application US/10273573
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: US/10/273,573
; PRIOR FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 8991
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (649)..(669)
; OTHER INFORMATION: CORONAVIRUS NUCLEOCAPSID PROTEIN domain identified by
; OTHER INFORMATION: EMATPIX, accession number DMO1206B, p-value=4.646e-09, raw score
US-10-273-573-8991

Query Match      26.1%; Score 47; DB 6; Length 1189;
Best Local Similarity 42.3%; Pred. No. 60;
Matches 11; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

QY      3 QSESSTPSTGCGSGKETPSRDDRSQS 28
DB      514 RSRFPSTTWSVMDTPQEKRPQS 539

RESULT 10
US-60-478-196-3144
; Sequence 3144, Application US/60478196
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Lemieux, Sebastien
; APPLICANT: Hu, Wengqi
; APPLICANT: Roemer, Terry
; TITLE OF INVENTION: IDENTIFICATION OF ESSENTIAL GENES OF ASPERGILLUS FUNIGATUS AND ME
; FILE REFERENCE: 10182-026-888
; CURRENT APPLICATION NUMBER: US/60/478,196
; PRIOR FILING DATE: 2003-06-13
; NUMBER OF SEQ ID NOS: 4000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3144
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-60-478-196-3144

Query Match      25.8%; Score 46.5; DB 7; Length 424;
Best Local Similarity 39.4%; Pred. No. 23;
Matches 13; Conservative 2; Mismatches 17; Indels 1; Gaps 1;

QY      3 QSESSTPSTGCGSGKETPSRDDRSQSRHHNGES 35
DB      17 QNPESTPSTPA-SKSSAPSTPSTEQSNMNAS 48

RESULT 11
PCT-US02-18638A-188
; Sequence 188, Application PC/TUS0218638A
; GENERAL INFORMATION:
```

```

; APPLICANT: Millennium Pharmaceuticals, Inc. et al.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; FILE REFERENCE: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035PC
; CURRENT APPLICATION NUMBER: PCT/US02/18638A
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 188
; LENGTH: 743
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-18638A-188

Query Match
Best Local Similarity 25.8%; Score 46.5; DB 1; Length 743;
Matches 13; Conservative 6; Mismatches 10; Indels 5; Gaps 1;

QY 1 SFQSSSTPSTGCGFSGKE-----TPEEDDRSQSR 29
DB 629 SRSSSSSSSTGSSSRDSSSTSSSSSSSRSR 662

RESULT 12
US-10-372-209-12
; Sequence 12, Application US/10372209
; GENERAL INFORMATION:
; APPLICANT: Hitachi, Ltd.
; TITLE OF INVENTION: RBMS-INTERACTING PROTEINS AND USE THEREOF
; FILE REFERENCE: HITA.0169
; CURRENT APPLICATION NUMBER: US/10/372,209
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: US 60/358,723
; PRIOR FILING DATE: 2002-02-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 186
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-372-209-12

Query Match
Best Local Similarity 25.6%; Score 46; DB 6; Length 186;
Matches 10; Conservative 7; Mismatches 11; Indels 2; Gaps 1;

QY 1 SFQSSSTPSTGCGFSGKETPEEDDRSQSR 30
DB 132 TFINNNSIPT--FDGEEBPSGQKVEOE 159

RESULT 13
US-10-372-209-13
; Sequence 13, Application US/10372209
; GENERAL INFORMATION:
; APPLICANT: Hitachi, Ltd.
; TITLE OF INVENTION: RBMS-INTERACTING PROTEINS AND USE THEREOF
; FILE REFERENCE: HITA.0169
; CURRENT APPLICATION NUMBER: US/10/372,209
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: US 60/358,723
; PRIOR FILING DATE: 2002-02-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 282
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-372-209-13
```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-372-209-13

Query Match
Best Local Similarity 25.6%; Score 46; DB 6; Length 282;
Matches 10; Conservative 7; Mismatches 11; Indels 2; Gaps 1;

QY 1 SFQSSSTPSTGCGFSGKETPEEDDRSQSR 30
DB 132 TFINNNSIPT--FDGEEBPSGQKVEOE 159

RESULT 14
US-10-372-209-18
; Sequence 18, Application US/10372209
; GENERAL INFORMATION:
; APPLICANT: Hitachi, Ltd.
; TITLE OF INVENTION: RBMS-INTERACTING PROTEINS AND USE THEREOF
; FILE REFERENCE: HITA.0169
; CURRENT APPLICATION NUMBER: US/10/372,209
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: US 60/358,723
; PRIOR FILING DATE: 2002-02-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 282
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-372-209-18

Query Match
Best Local Similarity 25.6%; Score 46; DB 6; Length 282;
Matches 10; Conservative 7; Mismatches 11; Indels 2; Gaps 1;

QY 1 SFQSSSTPSTGCGFSGKETPEEDDRSQSR 30
DB 132 TFINNNSIPT--FDGEEBPSGQKVEOE 159
```

```

US-10-380-565-34
; Sequence 34, Application US/10380565
; GENERAL INFORMATION:
; APPLICANT: Cooper, Joseph I
; TITLE OF INVENTION: Method and primers for detecting viral genes
; FILE REFERENCE: 0380-P0334US0
; CURRENT APPLICATION NUMBER: US/10/380,565
; CURRENT FILING DATE: 2003-06-17
; PRIOR APPLICATION NUMBER: PCT/GB01/04053
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: GB 0022505.2
; PRIOR FILING DATE: 2000-09-14
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 34
; LENGTH: 565
; TYPE: PRT
; ORGANISM: CABV
US-10-380-565-34

Query Match
Best Local Similarity 25.6%; Score 46; DB 6; Length 565;
Matches 14; Conservative 5; Mismatches 12; Indels 2; Gaps 1;

QY 2 FQSSSTPSTGCGFSGKETPEEDDRSQSRHNGE 34
DB 86 FNREASGSLGGQSHET--EVAFTQROEKFE 116
```

Search completed: July 24, 2003, 12:16:53

Fri Jul 25 15:44:28 2003

us-09-991-681-29.rapn

Page 5

Job time : 7.22489 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: July 24, 2003, 11:48:38 ; Search time 2.86969 Seconds
(without alignments)
1172.914 Million cell updates/sec

Title: US-09-991-681-29

Percent score: 180
Sequence: 1 SF0SESSTPTGSGKEPSEDDRSQSRHMGES 35

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: PIR_76:*
2: pir1:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	61	33.9	974	2 A40580	Iodestear maternal-
2	56	31.1	621	2 T48187	hypothetical prote
3	55	30.6	89	2 B95010	hypothetical prote
4	55	30.6	89	2 H97881	hypothetical prote
5	54.5	30.3	2514	2 F81045	hemagglutinin/hemo
6	54	30.0	715	2 D96582	hypothetical prote
7	53.5	29.7	737	1 K1RTCE	protein kinase C (
8	53.5	29.7	737	1 K1MSCE	protein kinase C (
9	53	29.4	384	2 T23604	hypothetical prote
10	53	29.4	390	2 G01936	Abi binding protei
11	53	29.4	814	2 T26702	hypothetical prote
12	53	29.4	3938	2 T42761	Baason protein -
13	52	28.9	218	2 C86337	hypothetical prote
14	52	28.9	675	2 S51037	zinc-finger protei
15	52	28.9	775	2 A61228	collagen alpha 2(I
16	52	28.9	976	1 TVMSMD	macrophage colony-
17	52	28.9	978	1 S16385	proliferating cell
18	52	28.9	2248	2 A35938	Carbon catabolite
19	51.5	28.6	427	2 JN0785	protein P21H1.2 (
20	51.5	28.6	2700	2 D88450	short neurotoxin 1
21	51	28.3	93	1 NNJ1F	HNS-type DNA bindi
22	51	28.3	92	2 AF3242	crank protein - Esc
23	51	28.3	132	2 S22997	transfer origin pr
24	51	28.3	132	2 T08535	hypothetical prote
25	51	28.3	210	2 F72645	hypothetical prote
26	51	28.3	414	2 T34000	hypothetical prote
27	51	28.3	610	2 S41315	hypothetical prote
28	51	28.3	642	2 C6152	T7123.2 protein -
29	51	28.3	1711	1 A47392	chromodomain-helic

30	51	28.3	1748	2 S42136	cnjB protein - Tet
31	50.5	28.1	430	2 T18651	hypothetical prote
32	50.5	28.1	522	2 S41819	nucleoporin p62 -
33	50.5	28.1	579	2 A84227	hypothetical prote
34	50	27.8	280	2 JC2358	melanoma antigen M
35	50	27.8	454	2 T29917	hypothetical prote
36	50	27.8	671	2 D72346	chemotaxis sensor
37	50	27.8	666	2 S70180	chemotaxis protein
38	50	27.8	710	2 T41586	hypothetical serin
39	50	27.8	735	2 T45059	hypothetical prote
40	50	27.8	819	2 T08745	probable RNA helic
41	50	27.8	829	2 T29372	hypothetical prote
42	50	27.8	930	2 D71617	SERA antigen/papai
43	50	27.8	1403	1 A47328	natural killer cel
44	50	27.8	1770	2 A71517	hypothetical prote
45	50	27.8	1912	2 T29088	vitellogenin I pre

ALIGNMENTS

RESULT 1
A40580 Iodestear maternal-effect protein - fruit fly (Drosophila melanogaster)
N:Alternate names: probable nucleoside triphosphate binding protein Iodestear
C:Species: Drosophila melanogaster
C>Date: 03-Aug-1992 #sequence_revision 03-Aug-1992 #text_change 21-Jul-2000
R:Accession: A40580, S19008
R:Girdham, C.H.; Glover, D.M.
Genes Dev. 5, 1786-1799, 1991
A:Title: Chromosome tangling and breakage at anaphase result from mutations in Iodestear,
A:Reference number: A40580; MUID:92009170; PMID:1916263
A:Accession: A40580
A:Molecule type: mRNA
A:Residues: 1-974 <GIR>
A:Cross-References: GB:X62629; NID:98187; PTDN:CAA44496.1; PID:98188
C:Genetics:
A:Gene: Iodestear
A:Cross-References: FlyBase:FBgn0002542

Query Match 33.9%; Score 61; DB 2; Length 974;
Best Local Similarity 51.9%; Pred. No. 7.2;
Matches 14; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Db 4 SESSSTPTGSGKEPSEDDRSQSR 30
39 SKSRPSSAGVVIDTQSEESQSS 65

RESULT 2
T48187 hypothetical protein F7A7.180 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
R:Accession: T48187
R:Bevan, M.; Terry, N.; Ardiles, W.; Buysbaert, C.; Daseville, R.; De Clerck, R.; De
ewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.
submitted to the Protein Sequence Database, March 2000
A:Reference number: Z24487
A:Accession: T48187
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-621 <BRV>
A:Cross-References: EMBL:AL61946
A:Experimental source: cultivar Columbia; BAC clone F7A7.
C:Genetics:
A:Map Position: 5
A:Insertion: 65/3; 263/3; 403/1; 469/3; 495/2; 529/2; 592/3
A>Note: F7A7.180

Query Match 31.1%; Score 56; DB 2; Length 621;
Best Local Similarity 38.7%; Pred. No. 20;
Matches 12; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

QY 1 SFQSSSTPTGCGFSGKETPSDDRSQSREH 31
 Db 475 SVEHKSSTIYAVGGYDCKEYLTAEPRDPREH 505

RESULT 3

B95010

Hypothetical protein SP0088 [imported] - Streptococcus pneumoniae (strain TIGR4)

C:Species: Streptococcus pneumoniae

C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001

C:Accession: B95010

R:Reteljin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heidon, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Kadane, D.; Holtzapfel, nson, T.; Hickey, E.K.; Holt, I.E.

Science 293, 498-506, 2001

A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison, A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.

A:Reference number: A95000; MUID:21357209; PMID:11463916

A:Accession: B95010

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-89 <KUR>

A:Cross-references: GB:AE005672; PIDN:AAK74275.1; PID:g14971554; GSPDB:GN00164; TIGR:SP4

C:Genetics:

A:Gene: SP0088

Query Match 30.6%; Score 55; DB 2; Length 89;
 Best Local Similarity 31.2%; Pred. No. 3.8;
 Matches 10; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

QY 3 QSESSTPTGCGFSGKETPSDDRSQSREHMG 34
 Db 55 ESQSATKQGGQTAKGTAGEDESNACQTOEISQ 86

RESULT 4

H97881

Hypothetical protein spt0080 [imported] - Streptococcus pneumoniae (strain R6)

C:Species: Streptococcus pneumoniae

C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 22-Oct-2001

C:Accession: H97881

R:Hoeklin, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; Dehoff, B.S.; Hee, R.; LeBlanc, D.J.; Lee, L.N.; Letkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M

Y, P.; Sun, P.M.; Winkler, M.E.

Y. Bacteriol. 183, 5709-5717, 2001

A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaekunas, S.R.; A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.

A:Reference number: A97872; MUID:21429245; PMID:11544234

A:Accession: H97881

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-89 <KUR>

A:Cross-references: GB:AE007317; PIDN:AAK98884.1; PID:g15457615; GSPDB:GN00174

C:Genetics:

A:Gene: spt0080

Query Match 30.6%; Score 55; DB 2; Length 89;
 Best Local Similarity 31.2%; Pred. No. 3.8;
 Matches 10; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

QY 3 QSESSTPTGCGFSGKETPSDDRSQSREHMG 34
 Db 55 ESQSATKQGGQTAKGTAGEDESNACQTOEISQ 86

RESULT 5

P81045

hemagglutinin/hemolysin-related protein NME1768 [imported] - Neisseria meningitidis (str

C:Species: Neisseria meningitidis

C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001

C:Accession: P81045

Query Match 30.6%; Score 54; DB 2; Length 715;
 Best Local Similarity 36.8%; Pred. No. 42;
 Matches 14; Conservative 4; Mismatches 16; Indels 4; Gaps 1;

R:Reteljin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A. Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.; ri, H.; Qin, H.; Vamathavan, U.; Gill, J.; Scariato, V.; Maignani, V.; Piazza, M.

Science 287, 1809-1815, 2000

A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rapapoli, R.; Ver

A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A:Reference number: A81000; MUID:20175755; PMID:10710307

A:Accession: P81045

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-2514 <TEP>

A:Cross-references: GB:AE002526; GB:AE002098; NID:g7227015; PIDN:AAK42109.1; PID:g722702.

A:Experimental source: serogroup B, strain MC58

C:Genetics:

A:Gene: NME1768

Query Match 30.3%; Score 54.5; DB 2; Length 2514;
 Best Local Similarity 36.4%; Pred. No. 1.3e+02;
 Matches 16; Conservative 9; Mismatches 8; Indels 11; Gaps 3;

QY 3 QSESSTPTGCGFSGKETPSDDRSQSREHMGES 35
 Db 1537 QSQETTYHQTKGSLMSAGIGFTIGSKTNTQENOSQNEHTGST 1580

RESULT 6

D96582

Hypothetical protein P1511.23 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: D96582

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federici, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Malt, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: D96582

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-715 <STO>

A:Cross-references: GB:AE005173; NID:g4587556; PIDN:AMD25787.1; GSPDB:GN00141

C:Genetics:

A:Gene: P1511.23

A:Map position: 1

Query Match 30.0%; Score 54; DB 2; Length 715;
 Best Local Similarity 36.8%; Pred. No. 42;
 Matches 14; Conservative 4; Mismatches 16; Indels 4; Gaps 1;

QY 1 SFQSSSTPTGCGFSGKETPSDDRSQSREHMG 34
 Db 40 SASSTSSPTTIGLSLFGSASVKSSSSSSSSHPSVGE 77

RESULT 7

X1RTCE

protein kinase C (EC 2.7.1.1-) epsilon - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 11-Jun-1999

C:Accession: B28163; B26408; S00216

R:Ono, Y.; Fujii, T.; Ogita, K.; Kikkawa, U.; Igarashi, K.; Nishizuka, Y.

J. Biol. Chem. 263, 6927-6932, 1988

A:Title: The structure, expression, and properties of additional members of the protein 1

A:Reference number: A92717; MUID:88198270; PMID:2834397

A:Accession: B28163

A:Molecule type: DNA

A:Residues: 1-737 <ONO>

```

A:Cross-references: GB:M1833; NID:g206182; PIDN:AAA41872.1; PID:g206183
R:Housey, G.M.; O'Brien, C.A.; Johnson, M.D.; Kirschmeier, P.; Weinstein, I.B.
Proc. Natl. Acad. Sci. U.S.A. 84, 1065-1069, 1987
A:Title: Isolation of cDNA clones encoding protein kinase C: evidence for a protein kinase
A:Reference number: A94145; MUID:87147193; PMID:3469647
A:Accession: B26408
A:Molecule type: mRNA
A:Residues: 397-447; GORGHLDREDFSGAET, 467, /LSNPTLLLPDGPPLRLRQ, 487-545, 'C', 547-636
A:Cross-references: GB:M15523; NID:g206192; PIDN:AAA41877.1; PID:g206193
C:Comment: Protein kinase C epsilon and epsilon' appear to be encoded by the same gene
C:Function:
A:Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonin
A>Note: activity is calcium-independent, phospholipid-dependent, and activated by diacyl
C:Superfamily: protein kinase C delta; protein kinase C zinc-binding repeat homology; pr
C:Keywords: alternative splicing; ATP; autophosphorylation; duplication; phorbol ester b
C
F:156-161/Region: pseudophosphorylation motif
F:170-220/Domain: protein kinase C zinc-binding repeat homology <K21>
F:243-252/Domain: protein kinase C zinc-binding repeat homology <K22>
F:406-668/Domain: protein kinase homology <KIN>
F:414-422/Region: protein kinase ATP-binding motif
F:170,201,204,220/Binding site: zinc (His, Cys, Cys) #status predicted
F:183,186,209,212/Binding site: zinc (Cys, Cys, His, Cys) #status predicted
F:243,273,276,292/Binding site: zinc (His, Cys, Cys, Cys) #status predicted
F:256,259,281,284/Binding site: zinc (Cys, Cys, His, Cys) #status predicted
F:437,456,532,534/Active site: Lys, Glu, Asp, Lys #status predicted
F:703,710/Binding site: phosphate (Thr) (covalent) (by autophosphorylation) #status pred
Query Match 29.7%; Score 53.5; DB 1; Length 737;
Best Local Similarity 48.0%; Pred. No. 50;
Matches 12; Conservative 5; Mismatches 3; Indels 5; Gaps 1;

Oy 4 SSSSTPSTGFGSGKETPSDDRSQS 28
Db 327 AESQPPASG-----NSPSEDDRSKS 346

RESULT 8
KIMSCB
protein kinase C (EC 2.7.1.1-) epsilon - mouse
C:Species: Mus musculus (house mouse)
C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 21-Nov-1997
C:Accession: S02270
R:Schaepp, D.; Parker, P.J.; Bristol, A.; Kriz, R.; Knopf, J.
FEBS Lett. 243, 351-357, 1989
A:Title: Unique substrate specificity and regulatory properties of PKC-epsilon: a ration
A:Reference number: S02270; MUID:89137541; PMID:2917656
A:Accession: S02270
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-737 <SCH>
C:Function:
A:Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonin
A>Note: activity is calcium-independent, phospholipid-dependent, and activated by diacyl
C:Superfamily: protein kinase C delta; protein kinase C zinc-binding repeat homology; pr
C:Keywords: ATP; autophosphorylation; duplication; phorbol ester binding; phospholipid b
F:156-161/Region: pseudophosphorylation motif
F:170-220/Domain: protein kinase C zinc-binding repeat homology <K21>
F:243-252/Domain: protein kinase C zinc-binding repeat homology <K22>
F:406-668/Domain: protein kinase homology <KIN>
F:414-422/Region: protein kinase ATP-binding motif
F:170,201,204,220/Binding site: zinc (His, Cys, Cys, Cys) #status predicted
F:183,186,209,212/Binding site: zinc (Cys, Cys, His, Cys) #status predicted
F:243,273,276,292/Binding site: zinc (His, Cys, Cys, Cys) #status predicted
F:256,259,281,284/Binding site: zinc (Cys, Cys, His, Cys) #status predicted
F:437,456,532,534/Active site: Lys, Glu, Asp, Lys #status predicted
F:703,710/Binding site: phosphate (Thr) (covalent) (by autophosphorylation) #status pred
Query Match 29.7%; Score 53.5; DB 1; Length 737;
Best Local Similarity 48.0%; Pred. No. 50;
Matches 12; Conservative 5; Mismatches 3; Indels 5; Gaps 1;

Oy 4 SSSSTPSTGFGSGKETPSDDRSQS 28

```

```

Db 327 AESQPPASG-----NSPSEDDRSKS 346

RESULT 9
T23604
hypothetical protein K11D2.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C:Accession: T23604; T26115
R:Kershaw, J.
submitted to the EMBL Data Library, November 1996
A:Reference number: 219768
A:Accession: T23604
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-384 <NLT>
A:Cross-references: EMBL:T28115; PIDN:CAE05558.1; GSPDB:GN00019; CESP:K11D2.4
A:Experimental source: clone K11D2
R:Lennard, N.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z20153
A:Accession: T26115
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-384 <WT2>
A:Cross-references: EMBL:T281137; PIDN:CAE03472.1; GSPDB:GN00019; CESP:K11D2.4
A:Experimental source: clone W02D9
C:Genetics:
A:Gene: CESP:K11D2.4
A:Map position: 1
A:Introns: 46/1; 103/1; 159/3; 225/1; 296/1; 354/3

Query Match 29.4%; Score 53; DB 2; Length 384;
Best Local Similarity 38.2%; Pred. No. 30;
Matches 13; Conservative 3; Mismatches 18; Indels 0; Gaps 0;

Oy 1 SFOSSSTPSTGFGSGKETPSDDRSQREHMG 34
Db 256 STSSAQQTPTMGCGGCGTTPSSASSSGSSGKCD 289

RESULT 10
G01936
Ab1 binding protein 3 - human
C:Species: Homo sapiens (man)
C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 17-Jul-1998
C:Accession: G01936
R:Ren, R.
submitted to the EMBL Data Library, July 1995
A:Reference number: G08875
A:Accession: G01936
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-390 <REN>
A:Cross-references: EMBL:U01089; NID:g987264; PID:g987265
C:Genetics:
A:Gene: Ab1BP3
C:Superfamily: SH3 homology
F:335-382/Domain: SH3 homology <SH3>
Query Match 29.4%; Score 53; DB 2; Length 390;
Best Local Similarity 45.5%; Pred. No. 31;
Matches 15; Conservative 2; Mismatches 14; Indels 2; Gaps 1;

Oy 1 SFOSSSTPSTGFGSGKETPSDDRSQREHMG 33
Db 142 SVNQNRNTYSTSGSSGSHPS--SRSSSRNSNG 172

RESULT 11
T26702
hypothetical protein Y38H6C.14 - Caenorhabditis elegans

```


Matches 15; Conservative 3; Mismatches 8; Indels 14; Gaps 2;
 Qy 9 PS-TGFFSGKETP-----SEDDRSQSRHNGE 34
 |||||
 Db 149 PSCTGCFRSPGPKGKGEPYALSEDDRDYRNGEICE 188
 |||||

Search completed: July 24, 2003, 11:58:05
 Job time : 5.86969 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 11:46:57 ; Search time 1.58858 Seconds

(without alignments)
1036.105 Million cell updates/sec

Title: US-09-991-681-29

Perfect score: 180

Sequence: 1 SFQSSSTPTGFGFGKTEPSEDDRSQSRHMGES 35

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Swissprot_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	61	33.9	1061	LDS DROME	P34739 drosophila
2	57.5	31.9	592	ODP2_DICDI	P34413 dictyostell
3	56	31.1	5596	MDN1_HUMAN	O9nu22 homo sapien
4	55.5	30.8	490	FXN3_HUMAN	O00409 homo sapien
5	55	30.6	585	YMO4_SCHPO	O9hgm7 schizosacch
6	54.5	30.3	568	DONS_DROME	O9vna8 drosophila
7	54.5	30.3	573	ETRA3_HUMAN	O99504 homo sapien
8	53.5	29.7	668	SG1_PIG	O99194 sus scrofa
9	53.5	29.7	737	KPCE_MOUSE	P16054 mus musculu
10	53.5	29.7	737	KPCE_MOUSE	P09216 rattus norv
11	52.5	29.2	508	NO60_DROME	O44081 drosophila
12	52	28.9	675	ZG20_KENIA	P18714 xenopus lae
13	52	28.9	977	KFMS_MOUSE	P09581 mus musculu
14	52	28.9	978	KFMS_MOUSE	O00495 rattus norv
15	52	28.9	981	UBPF_HUMAN	O944e8 homo sapien
16	52	28.9	981	UBPF_MOUSE	O845h1 mus musculu
17	51.5	28.6	427	CREA_ASPPG	O05620 aspergillus
18	51	28.3	83	CREA_NAAT	P01430 naia atra
19	51	28.3	131	TRK5_ECOLI	P17908 escherichia
20	51	28.3	613	DNAK_THEVO	O97b98 thermoplasm
21	51	28.3	757	TAU_HUMAN	P16036 homo sapien
22	51	28.3	1709	CHDI_HUMAN	O14646 homo sapien
23	51	28.3	1711	CHDI_MOUSE	P40201 mus musculu
24	50.5	28.1	522	NI62_HUMAN	P37198 homo sapien
25	50.5	28.1	969	AVVC_MOUSE	P96203 mus musculu
26	50	27.8	309	MAGI_HUMAN	P43355 homo sapien
27	50	27.8	671	CHEA_THENA	O56310 thermocoga
28	50	27.8	686	CHEA_RHOSH	O51135 rhodospacer
29	50	27.8	1462	NKCR_HUMAN	P34414 homo sapien
30	50	27.8	1770	PMPC_CHLTR	O84419 chlamydia t
31	50	27.8	1912	VITI_CHICK	P87498 gallus galli
32	49.5	27.5	168	VLPD_MYCHR	O49536 mycoplasma
33	49.5	27.5	515	STF7_YEAST	P06784 saccharomyc

34	49.5	27.5	736	KPCE_RABIT	P10830 oryctolagus
35	49.5	27.5	737	KPCE_HUMAN	O02156 homo sapien
36	49.5	27.5	3178	YS89_CAREL	O09624 caenorhabdi
37	49	27.2	278	MA32_MOUSE	O35658 mus musculu
38	49	27.2	278	MA32_RAT	O35786 mus musculu
39	49	27.2	326	2265_MOUSE	O91020 rattus norv
40	49	27.2	332	2265_RAT	O35968 rattus norv
41	49	27.2	821	GYRA_BACSU	P05653 bacillus su
42	49	27.2	851	CT06_HUMAN	O95501 homo sapien
43	49	27.2	1127	Y855_TREPA	O83827 treponema p
44	48.5	26.9	189	BCP_FEA	O41001 pisinu sativ
45	48.5	26.9	511	NF60_LOLPE	O01240 loligo peal

ALIGNMENTS

RESULT 1
LDS DROME STANDARD; PRT; 1061 AA.
AC P34739; O9VH1;
DT 01-FEB-1994 (Rel. 28, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Probable heliase lodestar.
GN LDS OR CG2684.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Preygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92009170; PubMed=1916263;
RA Girdham C.G., Glover D.M.;
RT "Chromosome tangling and breakage at anaphase result from mutations
RT in lodestar, a Drosophila gene encoding a putative nucleoside
RT triphosphate-binding protein.";
RL Genes Dev. 5:1786-1799(1991).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abriil J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Baau A., Bakendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Berns P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Broctler P.,
RA Butelis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Daventport L.B., Davies P.,
RA de Pallos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Dudin K.J., Engelstista C.C., Ferraz C., Ferrier S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodet K., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Honck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jatalaj M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Liako P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclob J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheefer J., Shen H.,
RA Shie B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

```

RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Massarman D.A., Weinstein G.M., Weisenbach J.,
RA Williams S.M., Moodie T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan W., Zhang S., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RT Science 287:2185-2195(2000).
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: BELONGS TO THE SNF2/RAD54 HELICASE FAMILY.
CC -!- CAUTION: Ref.1 sequence differs from that shown due to a
CC frameshift in position 946.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X62629; CAA44496.1; ALT_FRAME.
CC EMBL: AE003676; AAF54167.1; -.
CC Flybase: FBgn0002542; lds.
CC GO: GO:0008094; P:DNA dependent adenosinetriphosphatase activity; IDA.
CC GO: GO:0006353; P:transcription termination; IDA.
CC InterPro: IPR001410; DEAD.
CC InterPro: IPR001650; Helicase_C.
CC InterPro: IPR000330; SNF2_N.
CC Pfam: PF00271; helicase_C.1.
CC Pfam: PF00176; SNF2_N.1.
CC SMART: SM00487; DEXDC; 1.
CC SMART: SM00490; HELIC; 1.
CC Nuclear protein; Helicase; ATP-binding.
CC NP_BIND 465 472 ATP (POTENTIAL).
CC SITE 603 606 DEAD BOX.
CC FT SITE 167 167 R -> G (IN REF. 1).
CC FT CONFLICT 167 167
CC FT SEQUENCE 1061 AA; 118374 MW; 536BC893B1A90509 CRC64;
SQ
Query Match 33.9%; Score 61; DB 1; Length 1061;
Best Local Similarity 51.9%; Pred. No. 4.3;
Matches 14; Conservative 4; Mismatches 9; Indels 0; Gaps 0;
QY 4 SSSSTPTGSGFSGKEPTSEDDRSQSRE 30
DB 39 SKSRPSAGVVIDETSEESQSE 65
RESULT 2
ID OP2_DICDI STANDARD; PRT; 592 AA.
AC P36413;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Dihydrolipoamide acetyltransferase component of pyruvate dehydrogenase
DE complex, mitochondrial precursor (EC 2.3.1.12) (E2) (PDC-E2)
DE (Fragment).
DE DLAA.
OS Dictyostelium discoideum (slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_Taxid=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX2;
RL Mueller-Taubenberger A.;
RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: The pyruvate dehydrogenase complex catalyzes the overall
CC conversion of pyruvate to acetyl-CoA and CO(2). It contains
CC multiple copies of three enzymatic components: pyruvate
CC dehydrogenase (E1), dihydrolipoamide acetyltransferase (E2) and
CC lipoamide dehydrogenase (E3).
CC -!- CATALYTIC ACTIVITY: Acetyl-CoA + dihydrolipoamide = CoA + S-

```

```

CC acetyl-dihydrolipoamide.
CC -!- CORRELATOR: THE E2 COMPONENT CONTAINS TWO COVALENTLY-BOUND LIPOYL.
CC COFACTOR.
CC -!- SUBUNIT: 20 TO 30 ALPHA(2)-BETA(2) TETRAMERS OF E1 + 6 HOMODIMERS
CC OF E3 + 60 COPIES OF E2.
CC -!- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -!- SIMILARITY: BELONGS TO THE 2-OXOACID DEHYDROGENASE FAMILY.
CC -!- SIMILARITY: Contains 2 lipoyl-binding domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U06634; AAA1651.1; -.
CC HSP: P10515; IRYC.
CC Dictydb; DD05070; diaa.
CC InterPro: IPR001078; 2Oxoacid_dh.
CC InterPro: IPR006257; Acet_long.
CC InterPro: IPR000089; Biotin_lipoyl.
CC InterPro: IPR004167; E3 binding.
CC InterPro: IPR003016; Lipoyl.
CC Pfam: PF00198; 2-oxoacid_dh; 1.
CC Pfam: PF00364; biotin_lipoyl; 2.
CC Pfam: PF02817; e3_binding; 1.
CC ProDom: PD001115; 2Oxoacid_dh; 1.
CC TIGRFAMs: TIGR01349; PDHac_trf_mito; 1.
CC PROSITE: PS00189; LIPOYL; 2.
CC KW Glycolysis; Transferase; Acyltransferase; Repeat; Mitochondrion;
CC Transic peptide; Lipoyl.
CC FT NON_TER 1 1
CC FT TRANSIT 1 1
CC FT CHAIN ? 592
CC FT DOMAIN 35 130 MITOCHONDRION.
CC FT DOMAIN 158 241 DIHYDROLIPOAMIDE ACETYLTRANSFERASE
CC FT DOMAIN 258 271 COMPONENT OF PYRUVATE DEHYDROGENASE
CC FT DOMAIN 297 336 COMPLEX.
CC FT DOMAIN 360 592
CC FT BINDING 82 82 LIPOYL (BY SIMILARITY)..
CC FT BINDING 204 204 LIPOYL (BY SIMILARITY)..
CC FT BINDING 592 AA; 64424 MW; 30D0989F88D5BB7 CRC64;
SQ
Query Match 31.9%; Score 57.5; DB 1; Length 592;
Best Local Similarity 40.0%; Pred. No. 6.3;
Matches 14; Conservative 5; Mismatches 15; Indels 1; Gaps 1;
QY 1 SFQSSSTPTGSGFSGKE-TPSEDDRSQSREHME 34
DB 262 SSSSQSSTPTSSSSSSSQSTPTSSSQSTTRKSG 296
RESULT 3
ID MDN1_HUMAN STANDARD; PRT; 5596 AA.
AC Q9NU22; O15019;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Midasin (MIDAS-containing protein).
DE MDN1 OR KIAA0301.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RC PubMed=12102729;

```


RA Garbarino J.E., Gibbons I.R.;
 RT "Expression and genomic analysis of midasin, a novel and highly
 RT conserved AAA protein distantly related to dynein.";
 RL BMC Genomics 3:18-18(2002).
 RN [2]
 RP SEQUENCE OF 1255-2356 AND 3550-5596 FROM N.A.
 RC TISSUE=Testis;
 RA Tracey A.;
 RL Submitted (Apr-2000) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 3550-5596 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=97349984; PubMed=9205841;
 RA Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N.,
 RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. VII.
 RT The complete sequences of 100 new cDNA clones from brain which can
 RT code for large proteins in vitro.";
 RL DNA Res. 4:141-150(1997).
 CC -1- FUNCTION: May function as a nuclear chaperone and be involved in
 CC the assembly/disassembly of macromolecular complexes in the
 CC nucleus.
 CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -1- SIMILARITY: Contains 1 VMPA domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF503925; AAM77722.1; -;
 DR EMBL; AL096678; CAB86660.1; -;
 DR EMBL; AL096678; CAB86661.1; -;
 DR EMBL; AB002299; BAA20761.1; -;
 DR Genew; HGNC:18302; MDN1.
 DR GO; GO:0005634; C:nucleus; NAS.
 DR GO; GO:0016887; F:ATPase activity; NAS.
 DR GO; GO:0003754; F:chaperone activity; NAS.
 DR GO; GO:0006461; P:protein complex assembly; NAS.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR002035; VMP_A.
 DR Pfam; PF00004; AAA; 1.
 DR SMART; SM00382; AAA; 7.
 DR SMART; SM00327; VMA; 1.
 DR PROSITE; PS50234; VMPA; 1.
 KM Chaperone; ATP-binding; Repeat; Nuclear protein.
 FT NP_BIND 329 336 ATP (POTENTIAL).
 FT NP_BIND 677 684 ATP (POTENTIAL).
 FT NP_BIND 1084 1091 ATP (POTENTIAL).
 FT NP_BIND 1390 1397 ATP (POTENTIAL).
 FT NP_BIND 1753 1760 ATP (POTENTIAL).
 FT NP_BIND 2066 2073 ATP (POTENTIAL).
 FT DOMAIN 3566 3573 POLY-GLU.
 FT DOMAIN 4784 4791 POLY-GLU.
 FT DOMAIN 5008 5013 POLY-GLU.
 FT DOMAIN 5182 5187 POLY-GLU.
 FT DOMAIN 5384 5583 VMPA.
 FT CONFLICT 2287 2312 BLFSLMDPVHGDISRPMNRGLLEYI -> S (IN REF.
 FT 2).
 SQ SEQUENCE 5596 AA; 632802 MW; 586C62616A1F96D4 CRC64;
 Query Match 31.1%; Score 56; DB 1; Length 5596;
 Best Local Similarity 40.0%; Pred. No. 1.3e+02;
 Matches 12; Conservative 6; Mismatches 8; Indels 4; Gaps 1;
 Oy 3 QSESTPTGFGSKETPSDDRSQSRHM 32
 Db 4932 QNESQSPQ---EPREGSEDDKAGBEM 4957

RESULT 4
 FXN3 HUMAN STANDARD; PRT; 490 AA.
 ID FXN3_HUMAN
 AC 000409; 096117; 09UE17;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Checkpoint suppressor 1 (Forhead box protein N3).
 GN CHEK1 OR FOXN3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCB1_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=97299653; PubMed=9154802;
 RA Paci D., Keller C., Groudine M., Pion S.E.;
 RT "Reconstitution of a MEC1-independent checkpoint in yeast by
 RT expression of a novel human fork head cDNA.";
 RL Mol. Cell. Biol. 17:3037-3046(1997).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE=Muscle;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA DiCicco L., Marusik K., Farmer A.A., Rubin G.W., Hong L.,
 RA Stetson M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshitsuki S., Carrinchi P., Prange C.,
 RA Bask S.S., Loguallano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulys S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Wyer R.M.,
 RA Butterfield Y.S.N., Krzywicki M.I., Skalek U., Smalins D.E.,
 RA Schermer A., Schein J.B., Jones S.J.M., Maier M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP SEQUENCE OF 1-226 FROM N.A.
 RA Rowen L., Madan A., Qin S., Abbasi N., Baradaran L., Birditt B.,
 RA Bloom S., Dors M., Dichtoff R., Fleetwood P., Harrison G., James R.,
 RA Kaur A., Madan A., Owen M.P., Ratcliffe A., Shaffer T., Hood L.;
 RL Submitted (DEC-1999) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: PROBABLE TRANSCRIPTIONAL ACTIVATOR THAT MAY BE INVOLVED
 CC IN DNA DAMAGE-INDUCIBLE CELL CYCLE ARRESTS (CHECKPOINTS).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=000409-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=000409-2; Sequence=VSP_001551;
 CC Note=No experimental confirmation available;
 CC -1- SIMILARITY: Contains 1 fork-head domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U68723; AAB58252.1; -;
 DR EMBL; BC007506; AAH07506.1; -;
 DR EMBL; AC007263; AAF18259.1; -;

```

DR HSP: 063245; 2HFH.
DR TRANSFAC: T04204;
DR Genew; HGNC:1928; CHESTL.
DR MIM; 602628;
DR GO; GO:0003700; P:transcription factor activity; TAS.
DR GO; GO:0000077; P:DNA damage checkpoint; TAS.
DR GO; GO:0000085; P:G2 phase of mitotic cell cycle; TAS.
DR InterPro; IPR001766; TF_Fork_head.
DR Pfam; PF00250; Fork_head; 1.
DR PRINTS; PR00053; FORKHEAD.
DR ProDom; PD000425; TF_Fork_head; 1.
DR SMART; SM00339; FH; 1.
DR PROSITE; PS00657; FORK_HEAD_1; 1.
DR PROSITE; PS00658; FORK_HEAD_2; 1.
DR PROSITE; PS00659; FORK_HEAD_3; 1.
DR Transcription regulation; Activator; DNA-binding; Nuclear protein;
KW Alternative splicing.
FT DNA BIND 113 204
FT VARSP LIC 249 271
FT
FT
SQ SEQUENCE 490 AA; 53834 MW; EE07975A592C8C01 CRC64;
/FTid=VSP_001551.
Query Match
Best Local Similarity 30.8%; Score 55.5; DB 1; Length 490;
Matches 15; Conservative 5; Mismatches 11; Indels 7; Gaps 2;

Qy 5 ESSTPSTGGFSGKPTPS---EDRSQS---REHNGES 35
Db 375 EGSEGGSGSFRSHSPSDTEEDDRKHQKQKPEKSLDGS 412

RESULT 5
YR04_SCHPO
ID YR04_SCHPO STANDARD; PRT; 585 AA.
AC Q9HGM7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein C543.04 in chromosome II.
GN SPBC543.04.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OC NCBI_TaxID=4896;
XX
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros V., Peat N., Hayles J., Baker S., Baaham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Lechner S., McDonald S., McLean J.,
RA Mooney P., Mould S., Mungall K., Murphy L., Niblett D., Odeil C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Ruttenford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Grymptre B.,
RA Welljens I., Vanstreels E., Rieger M., Schaefer M., Meiller-Auer S.,
RA Gabell C., Fuchs M., Fritz C., Holzer E., Meestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Manbuth R., Purrelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Kocher M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revelante J.L., Moreno S., Armstrong J., Forburg S.L.,
RA Cernutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,

```

```

RA Shpakovski G.V., Usagery D., Bartell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -1 SIMILARITY: BELONGS TO THE UPF0171 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AL391604; CAC05246.1;
DR GeneDB; Spombe; SPBC543.04;
DR InterPro; IPR005365; UPF0171.
DR Pfam; PF03666; UPF0171; 1.
KW Hypothetical protein.
SQ SEQUENCE 585 AA; 66868 MW; 92F53DFB5AC1BE59 CRC64;

Query Match
Best Local Similarity 46.4%; Pred. No. 13;
Matches 13; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

Qy 5 ESSTPSTGGFSGKPTPS---EDRSQS---REHNGES 32
Db 143 ESSHPSTGSFEYKSSSKRSMSLFFHV 170

RESULT 6
DONS_DROME
ID DONS_DROME STANDARD; PRT; 568 AA.
AC Q9YNA8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Downstream of son gene protein homolog.
GN CG2669.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxID=7227;
XX
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers J.-H.C., Blazej R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durkin K.J., Evangelista C.C., Ferraz C., Ferrieri S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasner K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hoston D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kerchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasako P., Lei Y., Levitsky A.A., Li Z., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milhina N.V., Moritz C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,

```

```

RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svrtkac R., Tecor C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weissenbach J., Weissenbach U.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -1- SIMILARITY: BELONGS TO THE DONSON FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF003603; AAF52038.1; -
DR FlyBae; FBgn0037316; CG2869.
RW Hypothetical protein.
SQ SEQUENCE 568 AA; 63437 MW; A7E29D4C0AFD1480 CRC64;

Query Match 30.3%; Score 54.5; DB 1; Length 568;
Best Local Similarity 36.8%; Pred. No. 15;
Matches 14; Conservative 5; Mismatches 14; Indels 5; Gaps 1;

Oy 1 SFOSESTPTGCGFGKETP-----SEDDRSQSRHMG 33
Db 325 SFNESTTSLGPEAGEDAPPPAQEDDDDDMDLSELG 362

RESULT 7
ID EY33 HUMAN STANDARD: PRT; 573 AA.
AC Q99504; Q95463; Q99813;
DT 15-JUL-1998 (Rel. 36, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Eyes absent homolog 3.
GN EY33.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Embryo;
RA MEDLINE=97172972; PubMed=9020840;
RA Abdelhak S., Kalatzis V., Heilig R., Compain S., Samson D.,
RA Vincent C., Weil D., Crnaud C., Sahly I., Leibovici M.,
RA Bitner-Glindicz M., Francis M., Lacombe D., Vigneon J.,
RA Charachon R., Boven K., Bededer P., van Regemorter N.,
RA Weissenbach J., Petit C.;
RT "A human homologue of the Drosophila eyes absent gene underlies
RT brachio-oto-renal (BOR) syndrome and identifies a novel gene
RT family.";
RL Nat. Genet. 15:157-164(1997).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Retina;
RA MEDLINE=97202104; PubMed=9049631;
RA Zimmerman J.E., Bui Q.T., Steingrimson E., Nagle D.L., Fu W.,
RA Genin A., Spinner N.B., Copeland N.G., Jenkins N.A., Bucan M.,
RA Bonini N.M.;
RT "Cloning and characterization of two vertebrate homologs of the
RT Drosophila eyes absent gene";
RL Genome Res. 7:128-141(1997).
RN [3]

```

```

RP SEQUENCE OF 413-531 FROM N.A.
RX MEDLINE=99105912; PubMed=9887327;
RA Borsiani G., Degrandi A., Ballabio A., Bulfone A., Bernard L.,
RA Banfi S., Gattuso C., Mariani M., Dixon M., Donati D., Metcalfe K.,
RA Winter R., Robertson M., Axton R., Brown A., van Heyningen V.,
RA Hanson I.;
RT "EY44, a novel vertebrate gene related to Drosophila eyes absent.";
RL Hum. Mol. Genet. 8:11-23(1999).
CC -1- FUNCTION: MAY BE INVOLVED IN DEVELOPMENT OF THE EYE.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q99504-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q99504-2; Sequence=VSP 001493;
CC Note=No experimental confirmation available;
CC -1- SIMILARITY: BELONGS TO THE EYA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y10262; CAAT7311.1; -
DR EMBL; U81602; ABA42066.1; -
DR EMBL; AJ007991; CAA07814.1; -
DR Genew; HGNC:3521; EY33.
DR MIM; 601655; -
DR GO; GO:0007345; P:embryogenesis and morphogenesis; TAS.
DR GO; GO:0007601; P:vision; TAS.
DR InterPro; IPR006545; EYA.
DR InterPro; IPR005834; Hydrolyase.
DR Pfam; PF00702; Hydrolyase; 1.
DR TIGRFAMs; TIGR01658; EYA-cons domain; 1.
RW Developmental protein; Multigene family; Alternative splicing.
KV VARSPLIC 1 126
FT FT 142 142 /FTId=VSP_001493.
FT FT 151 151 H -> R (IN REF. 2).
FT FT 253 253 L -> V (IN REF. 2).
FT FT 268 268 K -> R (IN REF. 2).
FT FT 283 283 T -> S (IN REF. 2).
FT FT 290 290 N -> T (IN REF. 2).
FT FT 305 305 K -> R (IN REF. 2).
FT FT 305 305 V -> L (IN REF. 2).
FT FT 345 345 K -> E (IN REF. 2).
FT FT 357 357 S -> F (IN REF. 2).
FT FT 361 361 K -> E (IN REF. 2).
FT FT 373 374 PN -> SD (IN REF. 2).
FT FT 376 376 K -> N (IN REF. 2).
FT FT 379 379 N -> D (IN REF. 2).
FT FT 381 381 N -> S (IN REF. 2).
FT FT 388 388 R -> D (IN REF. 2).
FT FT 447 447 R -> K (IN REF. 1).
FT FT 449 449 R -> K (IN REF. 1).
FT FT 457 457 R -> N (IN REF. 1).
FT FT 473 473 R -> K (IN REF. 1).
FT FT 480 480 L -> P (IN REF. 2).
FT FT 487 487 L -> L (IN REF. 2).
FT FT 500 500 E -> K (IN REF. 1).
FT FT 524 526 SRF -> TSL (IN REF. 1).
FT FT 530 530 V -> L (IN REF. 1).
SQ SEQUENCE 573 AA; 62558 MW; EBE1F6435295CACS CRC64;

Query Match 30.3%; Score 54.5; DB 1; Length 573;
Best Local Similarity 31.9%; Pred. No. 15;
Matches 15; Conservative 6; Mismatches 11; Indels 15; Gaps 1;

Oy 1 SFOSE-----SPTPTGCGFGKETPSEDDRSQSRHMG 32
Db 236 TYOSEKPSVMAFAPAAQKSSGDPSTPSLSQTPTSKDTDDQSRKM 282

```


DR GO:0004699; F:calcium independent protein kinase C activity; IDA.
 DR InterPro: IPR000008; C2.
 DR InterPro: IPR002219; DAG_PE-bind.
 DR InterPro: IPR000961; PKinase_C.
 DR InterPro: IPR000719; Prot_Kinase.
 DR InterPro: IPR002290; Ser_thr_kinase.
 DR Pfam: PF00168; C2; 1.
 DR Pfam: PF00130; DAG_PE-bind; 2.
 DR Pfam: PF00069; PKinase; 1.
 DR Pfam: PF00433; PKinase_C; 1.
 DR PRINTS: PR00008; DAGPEDOMAIN.
 DR ProDom: PD000001; Prot_Kinase; 1.
 DR SMART: SM00109; C1; 2.
 DR SMART: SM00239; C2; 1.
 DR SMART: SM00133; S_TK_X; 1.
 DR SMART: SM00220; S_TKC; 1.
 DR PROSITE: PS50004; C2 DOMAIN 2; 1.
 DR PROSITE: PS00479; DAG_PE_BIND_DOM_1; 2.
 DR PROSITE: PS50081; DAG_PE_BIND_DOM_2; 2.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 KW Repeat: ATP-binding; Transferase; Phosphorylation;
 Setine/threonine-protein kinase; Phorbol-ester binding; zinc.
 FT DOMAIN 1 99
 FT DOMAIN 170 220 PHORBOL-ESTER AND DAG BINDING 1.
 FT DOMAIN 243 292 PHORBOL-ESTER AND DAG BINDING 2.
 FT DOMAIN 408 668 PROTEIN KINASE.
 FT NP_BIND 414 422 ATP (BY SIMILARITY).
 FT BINDING 437 437 ATP (BY SIMILARITY).
 FT ACT_SITE 532 532 BY SIMILARITY.
 FT MOD_RES 703 703 PHOSPHORYLATION (AUTO-) (POTENTIAL).
 FT MOD_RES 710 710 PHOSPHORYLATION (AUTO-) (POTENTIAL).
 SQ SEQUENCE 737 AA; 83560 MM; 7AAB8CC10C99F57 CRC64;

Query Match 29.7%; Score 53.5; DB 1; Length 737;
 Best Local Similarity 48.0%; Pred. No. 27;
 Matches 12; Conservative 5; Mismatches 3; Indels 5; Gaps 1;

Qy 4 SESSTPGSGFKETPSDDRSOS 28
 Db 327 AESPPQASG-----NSPSEDDRSKS 346

RESULT 10
 KPCE_RAT
 ID_KPCE_RAT STANDARD; PRT; 737 AA.
 AC P09216;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Protein kinase C, epsilon type (EC 2.7.1.-) (mKc-epsilon).
 GN PRKCE OR FKCE.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RC MEDLINE=88198270; PubMed=2834397;
 RA Ono Y., Fujii T., Ogita K., Kikkawa U., Igatahshi K., Nishizuka Y.;
 RT "The structure, expression, and properties of additional members of
 RT the protein kinase C family.";
 RL J. Biol. Chem. 263:6927-6932 (1988).
 RN [2]
 RP SEQUENCE OF 135-297 FROM N.A.
 RC MEDLINE=88083621; PubMed=3691811;
 RA Ono Y., Fujii T., Ogita K., Kikkawa U., Igatahshi K., Nishizuka Y.;
 RT "Identification of three additional members of rat protein kinase C
 RT family: delta, epsilon, and zeta-subspecies.";
 RL FEBS Lett. 226:125-128 (1987).
 CC -1- FUNCTION: THIS IS CALCIUM-INDEPENDENT, PHOSPHOLIPID-DEPENDENT,

CC SERINE- AND THREONINE-SPECIFIC ENZYME.
 CC -1- FUNCTION: PKC IS ACTIVATED BY DIACYLGLYCEROL WHICH IN TURN
 CC PHOSPHORYLATES A RANGE OF CELLULAR PROTEINS. PKC ALSO SERVES AS
 CC THE RECEPTOR FOR PHORBOL ESTERS. A CLASS OF TUMOR PROMOTERS.
 CC -1- SIMILARITY: Contains 2 zinc-dependent phorbol-ester and DAG
 CC binding domains.
 CC -1- SIMILARITY: Contains 1 C2 domain.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC PKC SUBFAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).

CC EMBL: M18331, AAA41872.1; -.
 CC PIR: B28163, KIRCE.
 CC PDB: 1GMI; 25-OCT-01.
 DR InterPro: IPR000008; C2.
 DR InterPro: IPR002219; DAG_PE-bind.
 DR InterPro: IPR000961; PKinase_C.
 DR InterPro: IPR000719; Prot_Kinase.
 DR InterPro: IPR002290; Ser_thr_kinase.
 DR Pfam: PF00168; C2; 1.
 DR Pfam: PF00130; DAG_PE-bind; 2.
 DR Pfam: PF00069; PKinase; 1.
 DR Pfam: PF00433; PKinase_C; 1.
 DR PRINTS: PR00008; DAGPEDOMAIN.
 DR ProDom: PD000001; Prot_Kinase; 1.
 DR SMART: SM00109; C1; 2.
 DR SMART: SM00239; C2; 1.
 DR SMART: SM00133; S_TK_X; 1.
 DR SMART: SM00220; S_TKC; 1.
 DR PROSITE: PS50004; C2 DOMAIN 2; 1.
 DR PROSITE: PS00479; DAG_PE_BIND_DOM_1; 2.
 DR PROSITE: PS50081; DAG_PE_BIND_DOM_2; 2.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 KW Repeat: ATP-binding; Transferase; Phosphorylation;
 Setine/threonine-protein kinase; Phorbol-ester binding; zinc;
 KW 3D-structure.
 FT DOMAIN 1 99
 FT DOMAIN 170 220 C2 DOMAIN.
 FT DOMAIN 243 292 PHORBOL-ESTER AND DAG BINDING 1.
 FT DOMAIN 408 668 PHORBOL-ESTER AND DAG BINDING 2.
 FT NP_BIND 414 422 PROTEIN KINASE.
 FT BINDING 437 437 ATP (BY SIMILARITY).
 FT ACT_SITE 532 532 ATP (BY SIMILARITY).
 FT MOD_RES 703 703 BY SIMILARITY.
 FT MOD_RES 710 710 PHOSPHORYLATION (AUTO-) (POTENTIAL).
 SQ SEQUENCE 737 AA; 83478 MM; 6AD6999EEDD2659F CRC64;

Query Match 29.7%; Score 53.5; DB 1; Length 737;
 Best Local Similarity 48.0%; Pred. No. 27;
 Matches 12; Conservative 5; Mismatches 3; Indels 5; Gaps 1;

Qy 4 SESSTPGSGFKETPSDDRSOS 28
 Db 327 AESPPQASG-----NSPSEDDRSKS 346

RESULT 11
 NOG0_DROME
 ID_NOG0_DROME STANDARD; PRT; 508 AA.
 AC O44081; O9V325;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Nucleolar protein at band 60B (minify protein).

GN N0608 OR MEL OR CG3333.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OK NCBI_Taxid=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9045289; PubMed=9829824;
 RA Phillips B., Billin A.N., Cadwell C., Buchholz R., Erickson C.,
 RA Merriam J.R., Carlson J., Poole S.J.;
 RT "The N0608 gene of Drosophila encodes an essential nuclear protein
 that functions in yeast."
 RL Mol. Genet. 260:20-29(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9189256; PubMed=10087258;
 RA Giordano E., Peluso I., Senger S., Furia M.;
 RT "Initially, a Drosophila gene required for ribosome biogenesis."
 RL J. Cell Biol. 144:1123-1133(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Gelinkner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,
 RA Abriil J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
 RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borrova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA de Pablos B., Delcher A., Deng Z., Davis A.D., Dew I., Dietz S.M.,
 RA Dodson K., Dong L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K., Eganellista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hoslin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kemison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Liao P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spirdling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-P., Zaveri J.S., Zhan W., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 CC -1- FUNCTION: PLAYS A CENTRAL ROLE IN RIBOSOMAL RNA PROCESSING.
 CC -1- SUBCELLULAR LOCATION: Nuclear; nucleolar.
 CC -1- SIMILARITY: BELONGS TO THE TRUB FAMILY OF PSEUDOURIDINE SYNTHASIS.
 CC -1- SIMILARITY: Contains 1 PUA domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch.
 CC -----
 CC EMBL: AF017230; AAC97117.1; -
 CC EMBL: AF089837; AAD16092.1; -
 CC EMBL: AE003463; AAF47178.1; -
 CC EMBL: FBgn0023184; N0608.
 CC InterPro: IPR004802; C055.
 CC InterPro: IPR002478; PUA.
 CC InterPro: IPR002501; TRUB_N.
 CC Pfam: PF01472; PUA; 1.
 CC Pfam: PF01509; TRUB_N; 1.
 CC SMART: SM00359; PUA; 1.
 CC TIGRFAMs: TIGR00425; C055; 1.
 CC PROSITE: PS50890; PUA; 1.
 CC Nuclear protein; tRNA processing.
 CC FT DOMAIN 294 369 PUA.
 CC FT DOMAIN 8 16 POLY-LYS.
 CC FT DOMAIN 450 453 POLY-ALA.
 CC FT DOMAIN 461 471 POLY-LYS.
 CC FT DOMAIN 489 499 POLY-LYS.
 CC SQ SEQUENCE 508 AA; 56830 MW; 3CAE3F91C84E0A94 CRC64;
 CC -----
 CC Query Match 29.2%; Score 52.5; DB 1; Length 508;
 CC Best Local Similarity 34.4%; Pred. No. 23;
 CC Matches 11; Conservative 8; Mismatches 12; Indels 1; Gaps 1;
 CC -----
 CC QY 4 SESSTPSTGSPS-GKTPSEDDRSQSRHMGCE 34
 CC DB 442 STSSVEETFAAASVSETPSKKKKKKKKKKD 473
 CC -----
 CC RESULT 12
 CC 2520_XENLA STANDARD; PRT; 675 AA.
 CC ID 2520_XENLA
 CC AC P18714;
 CC DT 01-NOV-1990 (Rel. 16, Created)
 CC DT 01-OCT-1996 (Rel. 34, Last sequence update)
 CC DT 15-SEP-2003 (Rel. 42, Last annotation update)
 CC DE Gastrula zinc finger protein XFG20-1 (XFG20.1).
 CC OS Xenopus laevis (African clawed frog).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 CC OC Xenopodinae; Xenopus.
 CC NCBI_Taxid=8355;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RX MEDLINE=95094815; PubMed=8001572;
 CC RT Schaefer U., Rausch O., Boumeester T., Pieler T.;
 CC "Sequence-specific recognition of a repetitive DNA element by a C2H2
 CC zinc-finger protein in Xenopus."
 CC RL Eur. J. Biochem. 226:567-576(1994).
 CC RN [2]
 CC RP SEQUENCE OF 85-613 FROM N.A.
 CC RX MEDLINE=90040698; PubMed=2509712;
 CC RA Nierfeld W., El-Baradi T., Mentzel H., Pieler T., Koester M.,
 CC RA Poeling A., Knoechel W.;
 CC "Second-order repeats in Xenopus laevis finger proteins."
 CC RL J. Mol. Biol. 208:639-659(1989).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.
 CC -----
 CC EMBL: X82643; CA57965.1; -
 CC DR PIR: S51037; S51037.
 CC DR HSP: P08046; IAI1.
 CC DR TRANSFAC: T02366; -
 CC DR InterPro: IPR007087; Znf_C2H2.
 CC DR Pfam: PF00096; Zf-C2H2; I8.

```

DR ProDom; PD000003; Znf_C2H2; 5.
DR SMART; SM00355; ZNF_C2H2; 18.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 17.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 18.
KW Zinc-finger; Metal-binding; DNA-binding; Repeat.
FT ZN_FING 62 84 C2H2-TYPE.
FT ZN_FING 90 112 C2H2-TYPE.
FT ZN_FING 118 140 C2H2-TYPE.
FT ZN_FING 146 168 C2H2-TYPE.
FT ZN_FING 174 196 C2H2-TYPE.
FT ZN_FING 202 224 C2H2-TYPE.
FT ZN_FING 257 279 C2H2-TYPE.
FT ZN_FING 286 308 C2H2-TYPE.
FT ZN_FING 344 366 C2H2-TYPE.
FT ZN_FING 373 395 C2H2-TYPE.
FT ZN_FING 424 446 C2H2-TYPE.
FT ZN_FING 452 474 C2H2-TYPE.
FT ZN_FING 507 529 C2H2-TYPE.
FT ZN_FING 535 557 C2H2-TYPE.
FT ZN_FING 563 585 C2H2-TYPE.
FT ZN_FING 591 613 C2H2-TYPE.
FT ZN_FING 619 642 C2H2-TYPE.
SQ SEQUENCE 675 AA; 77116 MW; 033094852C1FCF39 CRC64;

Query Match 28.9%; Score 52; DB 1; Length 675;
Best Local Similarity 37.5%; Pred. No. 38;
Matches 12; Conservative 8; Mismatches 10; Indels 2; Gaps 1;

OY 3 QSESTPSTGPGSGKETPSDDRSQREHMG 34
DB 392 QSTHTSPSTE--FGVQTEEDHMQSPSKDHTGE 421

RESULT 13
KEMS_MOUSE
ID KEMS_MOUSE STANDARD; PRT; 977 AA.
AC P09581; Q9DBH9;
DT 01-MAR-1989 (Rel. 10, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Macrophage colony stimulating factor I receptor precursor (CSF-1-R)
DE (EC 2.7.1.112) (fms proto-oncogene) (c-fms).
GN CSF1R OR CSFMR OR FMS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eubacteria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88217329; PubMed=2966922;
RA Rochwell V.M.; Rohrschneider L.R.;
RT "Murine c-fms cDNA: cloning, sequence analysis and retroviral
RT expression."; 1:311-324(1987).
RL Oncogene Res. 1:311-324(1987).
RN [2]
RP REVISIONS.
RA Rochwell V.M.;
RL Submitted (SEP-1988) to the EMBL/GenBank/DBJ databases.
RN [3]
RP REVISIONS.
RX MEDLINE=93181280; PubMed=8441691;
RA de Paraveval N.; Borderaux D.; Gieselbrecht S.; Sola B.;
RT "Reassessment of the murine c-fms proto-oncogene sequence.";
RL Nucleic Acids Res. 21:750-750(1993).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Liver;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J.; Shingawa A.; Shibata K.; Yoshino M.; Itoh M.; Ishi Y.;
RA Arikawa T.; Hara A.; Fukunishi Y.; Konno H.; Adachi U.; Fukuda S.;
RA Aizawa K.; Izawa M.; Nishi K.; Kiyosawa H.; Kondo S.; Yamada I.;
RA Saito T.; Okasaki Y.; Gojohori T.; Bono H.; Kasukawa T.; Saito R.;
RA Kadota K.; Matsuda H.A.; Ashburner M.; Batalov S.; Casavant T.;

```

```

RA Fleischmann W.; Gaasterland T.; Gissi C.; King B.; Kochiwa H.;
RA Kuehl P.; Lewis S.; Matsuo Y.; Nikaide I.; Pesole G.; Quackenbush J.;
RA Schiraldi L.M.; Staabli F.; Suzuki R.; Tomita M.; Wagner L.; Washio T.;
RA Sakai K.; Okido T.; Furuno M.; Aono H.; Baldarelli R.; Barab G.;
RA Blake J.; Botelli D.; Bojunga N.; Carninci P.; de Bona M.F.;
RA Brownstein M.J.; Bult C.; Fletcher C.; Fujita M.; Gariboldi M.;
RA Gustincich S.; Hill D.; Hofmann M.; Hume D.A.; Kamita M.; Lee N.H.;
RA Lyons P.; Marchionni L.; Mashima J.; Mazzarelli J.; Mombaerts P.;
RA Nordone P.; Ring B.; Ringwald M.; Rodriguez I.; Sakamoto N.;
RA Sasaki H.; Sato K.; Schoenbach C.; Seta T.; Shibata Y.; Storch K.-F.;
RA Suzuki H.; Toyooka K.; Wang K.H.; Weitz C.; Whitaker C.; Wilming L.;
RA Wynshaw-Boris A.; Yoshida K.; Hasegawa Y.; Kawai H.; Kohleuk S.;
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [5]
RP SEQUENCE OF 1-16 FROM N.A.
RX MEDLINE=93268269; PubMed=8497248;
RA Yue X.; Favot P.; Dunn T.L.; Casasady A.I.; Hume D.A.;
RT "Expression of mRNA encoding the macrophage colony-stimulating factor
RT receptor (c-fms) is controlled by a constitutive promoter and tissue-
RT specific transcription elongation.";
RL Mol. Cell. Biol. 13:3191-3201(1993).
RN [6]
RP AUTOPHOSPHORYLATION SITES.
RX MEDLINE=90258890; PubMed=2160591;
RA van der Geer P.; Hunter T.;
RT "Identification of tyrosine 706 in the kinase insert as the major
RT colony-stimulating factor I (CSF-1)-stimulated autophosphorylation
RT site in the CSF-1 receptor in a murine macrophage cell line.";
RL Mol. Cell. Biol. 10:2991-3002(1990).
CC -1- FUNCTION: THIS PROTEIN IS THE RECEPTOR FOR CSF-1. IT IS A PROTEIN
CC TYROSINE-KINASE TRANSMEMBRANE RECEPTOR.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -1- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; X06368; CAA29666.1; ALT SEQ.
DR EMBL; AK004947; BAB33691.1; -
DR EMBL; S62219; -; NOT_ANNOTATED_CDS.
DR HSSP; P11362; IFGK.
DR MGD; MGI:1339758; Cef1r.
DR InterPro; IPR007110; IG_1like.
DR InterPro; IPR003598; IG_C2.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001824; RTKinaseIII.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00047; IG_4.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 2.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00219; TyKc; 1.
DR PROSITE; PS50835; IG_LIKE; 4.
DR PROSITE; PS50107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
KW Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;
KW Transferrase; Glycoprotein; Phosphorylation; ATP-binding; Signal;
KW Immunoglobulin domain; Repeat.
FT SIGNAL 1 19 POTENTIAL.

```


RESULT 15
 ID UBP_HUMAN STANDARD; PRT; 981 AA.
 AC O9Y4E8; O9HCA6; O9UNP0; O9YSB5;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Ubiquitin carboxyl-terminal hydrolase 15 (EC 3.1.2.15) (Ubiquitin
 thiolesterase 15) (ubiquitin-specific processing protease 15)
 DE (Deubiquitinating enzyme 15) (Unph-2) (Unph4).
 GN USP15 OR KIAA0529.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE=Brain;
 RX MEDLINE=22158633; PubMed=12168954;
 RA Nakajima D., Okazaki N., Yamakawa H., Kikuno R., Ohara O., Nagase T.;
 RT "Construction of expression-ready cDNA clones for KIA genes: manual
 curation of 330 KIA cDNA clones.";
 RL DNA Res. 9:99-106(2002).
 RN [2]
 RP SEQUENCE OF 4-981 FROM N.A. (ISOFORM 2).
 RC TISSUE=Brain;
 RX MEDLINE=98290545; PubMed=9628581;
 RA Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,
 RA Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. IX.
 The complete sequences of 100 new cDNA clones from brain which can
 code for large proteins in vitro.";
 RL DNA Res. 5:31-39(1998).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE=Brain;
 RA Kim K.I., Nagase T., Chung C.H.;
 RT "Identification and characterization of a new human deubiquitinating
 enzyme Unph4.";
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Dermal papilla;
 RA Kim M.K., Kim Y.H., Seo J.M., Lee H.M., Chung H.J., Sohn M.Y.,
 RA Hwang S.Y., Im S.U., Jung E.J., Lee J.H., Kim J.C.;
 RT "A catalogue of genes in the human dermal papilla cells as identified
 by expressed sequence tags.";
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE OF 19-981 FROM N.A. (ISOFORM 3).
 RC TISSUE=Fetal brain;
 RA Kimura Y., Saya H., Nakao M.;
 RT "Cloning and identification of human Unph-2.";
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP IDENTIFICATION (ISOFORM 2), AND ACTIVITY.
 RX MEDLINE=99375319; PubMed=10444327;
 RA Baker R.T., Wang X.-W., Woolliatt E., White J.A., Sutherland G.R.;
 RT "Identification, functional characterization, and chromosomal
 localization of USP15, a novel human ubiquitin-specific protease
 related to the UNP oncoprotein, and a systematic nomenclature for
 human ubiquitin-specific proteases.";
 RL Genomics 59:264-274(1999).
 RN [7]
 RP ALTERNATIVE SPLICING (ISOFORMS 1 AND 2).
 RX MEDLINE=22419901; PubMed=12532266;
 RA Angelats C., Wang X.-W., Verma L.S., Copeland N.G., Jenkins N.A.,
 RA Baker R.T.;
 RT "Isolation and characterization of the mouse ubiquitin-specific
 protease Usp15.";
 RL Mamm. Genome 14:31-46(2003).
 CC -1- CATALYTIC ACTIVITY: Ubiquitin C-terminal thiolester + H(2)O =

CC ubiquitin + a thiol.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=3;
 CC Name=1;
 CC IsoId=Q9Y4E8-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q9Y4E8-2; Sequence=VSP_005261;
 CC Name=3;
 CC IsoId=Q9Y4E8-3; Sequence=VSP_005260;
 CC -1- TISSUE SPECIFICITY: Expressed in skeletal muscle, kidney, heart,
 CC placenta, liver, thymus, lung, and ovary, with little or no
 CC expression in other tissues.
 CC -1- SIMILARITY: Belongs to peptidase family C19.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 CC EMBL; AB011101; BA25455.2; -;
 CC EMBL; AF106069; AA052099.1; -;
 CC EMBL; AF153604; AA041086.1; -;
 CC EMBL; AF013990; AAC28973.1; -;
 CC MEROPS; C19_022; -;
 CC GeneW; HGNC:12613; USP15.
 CC MIM; 604731; -;
 CC GO; GO:0004197; F:cytosine-type endopeptidase activity; TAS.
 CC GO; GO:0004843; F:ubiquitin-specific protease activity; TAS.
 CC InterPro; IPR00615; DUSP.
 CC InterPro; IPR001394; UCH-2.
 CC Pfam; PF00443; UCH; 1.
 CC SMART; SM00695; DUSP; 1.
 CC PROSITE; PS00972; UCH_2_1; 1.
 CC PROSITE; PS00973; UCH_2_2; 1.
 CC PROSITE; PS0235; UCH_2_3; 1.
 CC Ubi conjugation pathway; Hydrolyase; Thiol protease; Multigene family;
 KW Alternative splicing.
 FT ACT_SITE 298 298 BY SIMILARITY.
 FT ACT_SITE 883 883 BY SIMILARITY.
 FT ACT_SITE 891 891 BY SIMILARITY.
 FT VARSPLIC 217 256 DGTWPGSTPKSPGASNFSTLPKISSLSNNYNNNNR
 FT -> OKNEDGTWPGSTPK (in isoform 3).
 FT FTID=VSP_005260.
 FT Missing (in isoform 2).
 FT VARSPLIC 228 256
 FT CONFLICT 559 559 T -> A (IN REF. 5).
 FT CONFLICT 968 968 N -> H (IN REF. 5).
 FT SEQUENCE 981 AA; 112418 MW; E81FEB9DE57F7069 CRC64;
 Query Match 28.9%; Score 52; DB 1; Length 981;
 Best Local Similarity 32.3%; Pred. No. 57;
 Matches 10; Conservative 7; Mismatches 14; Indels 0; Gaps 0;
 Cy 3 QSESSTGTGFGSKTSPEDDSQSRHMG 33
 Db 655 EMTDEPDDESSQDQELPSENNQSDSVG 685
 Search completed: July 24, 2003, 11:54:23
 Job time : 3.58856 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: July 24, 2003, 11:47:53 ; Search time 7.78917 Seconds

(without alignments)
1159.539 Million cell updates/sec

Title: US-09-991-681-29

Perfect score: 180

Sequence: 1 SFQSSSTPTGFGSGKETPSEDDRSQSRHMGES 35

Scoring table: BLOSUM62

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_23.*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp Vertebrate:*
14: sp Unclassified:*
15: sp Virus:*
16: sp Bacteriaph:*
17: sp_Archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	180	100.0	589	4 Q8N4Y4	Q8N4Y4 homo sapien
2	180	100.0	592	4 Q96CH9	Q96CH9 homo sapien
3	180	100.0	1770	4 Q9ULH6	Q9ULH6 homo sapien
4	61	33.9	1061	5 Q9VHY1	Q9VHY1 drosophila
5	59.5	33.1	1252	5 Q9BLU9	Q9BLU9 leishmania
6	59	32.8	1403	5 Q9VPO4	Q9VPO4 drosophila
7	58	32.2	252	11 Q9CKS4	Q9CKS4 mus musculu
8	57	31.7	118	3 Q96TV3	Q96TV3 pleurotus o
9	57	31.7	345	5 Q9VJ23	Q9VJ23 drosophila
10	56	31.1	621	10 Q9M010	Q9M010 arabidopsis
11	55	30.6	89	16 Q97T67	Q97T67 streptococc
12	55	30.6	89	16 Q8CZC2	Q8CZC2 streptococc
13	55	30.6	287	4 Q8IWH6	Q8IWH6 homo sapien
14	55	30.6	656	5 Q96J32	Q96J32 drosophila
15	54.5	30.3	536	4 Q8IVX7	Q8IVX7 homo sapien
16	54.5	30.3	2514	16 Q9UY30	Q9UY30 neisseria m

17	54	30.0	322	4 Q75370	Q75370 homo sapien
18	54	30.0	325	6 Q8HY94	Q8HY94 piecotus to
19	54	30.0	337	13 Q91708	Q91708 xenopus lae
20	54	30.0	351	16 Q8POE5	Q8POE5 streptococc
21	54	30.0	351	16 Q99Z56	Q99Z56 streptococc
22	54	30.0	525	12 Q83696	Q83696 measles vir
23	54	30.0	600	5 Q9VE96	Q9VE96 drosophila
24	54	30.0	712	10 Q9M5P5	Q9M5P5 arabidopsis
25	54	30.0	715	10 Q9SVH1	Q9SVH1 arabidopsis
26	54	30.0	956	4 Q9UQ39	Q9UQ39 homo sapien
27	54	30.0	1262	4 Q9UQ40	Q9UQ40 homo sapien
28	54	30.0	1808	13 Q42142	Q42142 gallus gall
29	54	30.0	2296	4 Q9UHA8	Q9UHA8 homo sapien
30	54	30.0	2607	11 Q8BRT8	Q8BRT8 mus musculu
31	54	30.0	2752	4 Q9UQ35	Q9UQ35 homo sapien
32	53.5	29.7	280	3 Q8X0J7	Q8X0J7 neurospora
33	53.5	29.7	1209	11 Q9CUE8	Q9CUE8 mus musculu
34	53.5	29.7	1297	4 Q150Z5	Q150Z5 homo sapien
35	53	29.4	384	5 Q02283	Q02283 caenorhabdi
36	53	29.4	390	4 Q13249	Q13249 homo sapien
37	53	29.4	614	5 P91956	P91956 lytechinus
38	53	29.4	814	5 Q9XX47	Q9XX47 caenorhabdi
39	53	29.4	3938	11 Q88778	Q88778 rattus norv
40	52.5	29.2	169	4 Q961I9	Q961I9 homo sapien
41	52.5	29.2	229	10 Q8S0B1	Q8S0B1 oryza sativ
42	52.5	29.2	290	4 Q8LAM7	Q8LAM7 arabidopsis
43	52.5	29.2	300	4 Q94871	Q94871 homo sapien
44	52.5	29.2	319	10 Q9C5J5	Q9C5J5 arabidopsis
45	52.5	29.2	363	10 Q9LMC6	Q9LMC6 arabidopsis

ALIGNMENTS

RESULT 1
ID Q8N4Y4 PRELIMINARY; PRT; 589 AA.
AC Q8N4Y4;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Similar to KIAA1244 protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strauberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL, BC033191; AAH33191.1; -.
FT NON_TER
SQ SEQUENCE 589 AA; 66086 MW; 8041EBA348DE65F7 CRC64;
Query Match 100.0%; Score 180; DB 4; Length 589;
Best Local Similarity 100.0%; Pred. No. 5.9e-17;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SFQSSSTPTGFGSGKETPSEDDRSQSRHMGES 35
Db 354 SFQSSSTPTGFGSGKETPSEDDRSQSRHMGES 388
RESULT 2
ID Q96CH9 PRELIMINARY; PRT; 592 AA.
AC Q96CH9;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strausberg R.;
RA Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC014227; AAH14227.1; -.
KW Hypothetical protein.
FT NON TER
SQ SEQUENCE 592 AA; 66400 MW; FA1E807BD0F47B5 CRC64;

Query Match 100.0%; Score 180; DB 4; Length 592;
Best Local Similarity 100.0%; Pred. NO. 5.9e-17;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFQSESTPTGTGFGSKETPSEDNRQSRHMGES 35
DB 357 SFQSESTPTGTGFGSKETPSEDNRQSRHMGES 391

RESULT 3
OY 09UH6 PRELIMINARY; PRT; 1770 AA.
AC 09UH6; 096P46; (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 22, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE B1G3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RA Hong W.;
RT "X1A1244 as a novel distantly related member (B1G3) of the B1G1/Sec7p
RL subfamily of ARF GEFs.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF413080; AL04174.1; -.
DR InterPro; IPR000904; Sec7.
DR SMART; SM00222; Sec7; 1.
KW Hypothetical protein.
SQ SEQUENCE 1770 AA; 195845 MW; SE99E36A6P92AB4 CRC64;

Query Match 100.0%; Score 180; DB 4; Length 1770;
Best Local Similarity 100.0%; Pred. NO. 2e-16;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFQSESTPTGTGFGSKETPSEDNRQSRHMGES 35
DB 1535 SFQSESTPTGTGFGSKETPSEDNRQSRHMGES 1569

RESULT 4
OY 09VH1 PRELIMINARY; PRT; 1061 AA.
AC 09VH1;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE CG2684 protein.
LDS OR CG2684.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;

```

```

RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Suton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Change M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Baas A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.A., Boulton J., Broststein P., Brotler P.,
RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA de Pablo S., Dopp L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibbegan C.,
RA Jaisli M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusken D.R., Paclab J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Part V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun B.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Sheng H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003676; AAF54167.1; -.
DR Flybase; FBgn0003542; lds.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR003330; SNF2_N.
DR Pfam; PF00271; helicase_C; 1.
DR Pfam; PF00176; SNF2_N; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
KW ATP-binding; Helicase; Hydrolase.
SQ SEQUENCE 1061 AA; 118374 MW; 536HC93B1A90509 CRC64;

Query Match 33.9%; Score 61; DB 5; Length 1061;
Best Local Similarity 51.9%; Pred. NO. 7.8;
Matches 14; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

OY 4 SSSSTPTGTGFGSKETPSEDNRQSRHMGES 30
DB 39 SSSSTPTGTGFGSKETPSEDNRQSRHMGES 65

RESULT 5
OY 09BLU9 PRELIMINARY; PRT; 1252 AA.
AC 09BLU9;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Possible SNF2-related helicase.
L2230.06.
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN (1)
RP SEQUENCE FROM N.A.

```

RC STRAIN=Friedlin;
 RA Zimmermann W., Wamburt R., Ivens A.C., Quail M., Rajandream M.A.,
 RA Barrell B.G.;
 RL Submitted (JAN-2001) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Friedlin;
 RX MEDLINE=98146435; PubMed=9477341;
 RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
 RA Smith D.F.;
 RT "A physical map of the leishmania major Friedlin genome.";
 RL Genome Res. 8:135-145(1998).
 DR EMBL; AL513062; CAC24685.1; -
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR001650; Helicase_C.
 DR InterPro; IPR000330; SNF2_N.
 DR Pfam; PF00271; helicase_C_1.
 DR Pfam; PF00176; SNF2_N; 1.
 DR SMART; SM00487; DEXDC; 1.
 DR SMART; SM00490; HELIC; 1.
 KW ATP-binding; Helicase; Hydrolase.
 SQ SEQUENCE 1252 AA; 137895 MW; 4E6D2927A8B1A745 CRC64;

Query Match 33.1%; Score 59.5; DB 5; Length 1252;
 Best Local Similarity 41.0%; Pred. No. 15;
 Matches 16; Conservative 4; Mismatches 12; Indels 7; Gaps 2;

OY 4 SSSSTPSTGTF---SGKET-----SEDDRSQSRHHMGS 35
 Db 395 SRASTQLTGTGTTGSRGKATPSTSLSSQOQOQKREAGSA 433

RESULT 6

O9VPO4 PRELIMINARY; PRT; 1403 AA.

AC O9VPO4
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE CG4297 protein.

OS Drosophila melanogaster (Fruit Fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoikins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Chame M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Barker E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abri J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktoglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Botkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Butts K.C., Busam D.A., Butler H., Cadieu L.B., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delecher A., Deng Z., Deyanov A., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Dudin K.J., Evangelista C.C., Ferrar C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Galbraith W.M., Glasser K.,
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris K.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hootin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
 RA Jaislin M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kilp D., Lai Z.,
 RA Laško P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Noshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusken D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Schaefer F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun B.,
 RA Svrtkars R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 DR EMBL; AB003589; AAF51491.1; -
 DR FLYBase; FBgn001258; CG4297.
 SQ SEQUENCE 1403 AA; 150676 MW; D7CE38B1DDC4ACDF CRC64;

Query Match 32.8%; Score 59; DB 5; Length 1403;
 Best Local Similarity 38.1%; Pred. No. 21;
 Matches 16; Conservative 5; Mismatches 13; Indels 8; Gaps 1;

OY 1 SPOSSSTPSTGTFSGKET-----PSEDDRSQSRHHMGE 34
 Db 1104 SSSSSSTSTSLSSLAGESDPNELALALEKEISQGTGEHMG 1145

RESULT 7

O9CKX4 PRELIMINARY; PRT; 252 AA.

AC O9CKX4
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE 3110013H01RIK protein.

OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Embryo;
 RX MEDLINE=21085660; PubMed=11217851;

RA Kawai T., Shinagawa A., Shibata K., Yoshino M., Itoh M., Iehi Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Komoto H., Adachi U., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
 RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gisi C., King B., Kochwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Resole G., Quackenbush J.,
 RA Schmitt L.M., Staubli F., Suzuki R., Tomita M., Wagner U., Washio T.,
 RA Blake J., Boffelli D., Fujuno M., Aono H., Baldarelli R., Barsh G.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gusticich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima M., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Welter C., Whitaker C., Wilming L.,
 RA Wyszewski B., Yoshida K., Hasegawa Y., Kawaji H., Kohatsu S.,
 RA Wyszewski B., Yoshida K., Hasegawa Y., Kawaji H., Kohatsu S.,
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 DR EMBL; AK014046; BA829130.1; -
 DR MGD; MGI:1920389; 3110013H01RIK.
 DR InterPro; IPR006913; DUF636.
 DR Pfam; PF04828; DUF636; 1.
 SQ SEQUENCE 252 AA; 27482 MW; FDEADB1DE2A69961 CRC64;

Query Match 32.2%; Score 58; DB 11; Length 252;
 Best Local Similarity 45.8%; Pred. No. 4.2;

Matches 11; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 10 STGSPGSGKTPSEDDRSQSRHMG 33
 DB 20 ATGGLSGSGSPQRSRSRTAGG 43

RESULT 8

Q96TV3 PRELIMINARY; PRT; 118 AA.

AC Q96TV3; 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 GN Hypothetical zinc finger protein (Fragment).
 OS Pleurotus ostreatus (Oyster mushroom) (White-rot fungus).
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
 OC Agaricales; Pleurotaceae; Pleurotus.
 OK NCBI_TaxID=5322;

Q96TV3 SEQUENCE FROM N.A.
 RA Pisabarro A.G.;
 RN Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
 RP [2]
 RA "Analays and genetic mapping of genes expressed in the lamellae of
 RT the edible basidiomycete Pleurotus ostreatus.";
 DR EMBL; AJ11658; CAD10796.1; -
 DR InterPro; IPR002328; ADH_zinc.
 DR InterPro; IPR002893; ZnF_MIND.
 DR Pfam; PF01753; zf-MIND; 1.
 DR PROSITE; PS00059; ADH_ZINC; 1.
 FT NON TER 1 118
 FT SEQUENCE 118 AA; 13189 MW; E25005B7C57BF772 CRC64;

Query Match 31.7%; Score 57; DB 3; Length 118;
 Best Local Similarity 42.4%; Pred. No. 2.5;
 Matches 14; Conservative 3; Mismatches 16; Indels 0; Gaps 0;

QY 2 FSESSTPSTGSGKTPSEDDRSQSRHMG 34
 DB 44 FSESASSISSTGSSVSSSEDRTSISGHEG 76

RESULT 9

Q9VJ23 PRELIMINARY; PRT; 345 AA.

AC Q9VJ23; 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE CG15169 protein.
 GN CG15169.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OK NCBI_TaxID=7227;

Q9VJ23 SEQUENCE FROM N.A.
 RA STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Glinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amaralides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vardell M.D., Zhang Q., Chen L.X.,
 RA Barton R.C., Rogers Y.-H.C., Blaise R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Planckoch C., Baldwin D.,

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
 RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K., Evangelista C.C., Ferraz C., Periera S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
 RA Jalali M., Kaush F., Karpen G.H., Ke Z., Kennison J.A., Kerchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milehina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
 RA Palazzolo M., Piltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Splier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J., Ye J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu Z., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 DR EMBL; AB03661; AA53734.1;
 DR Flybase; FBgn032734; CG15169.
 SQ SEQUENCE 345 AA; 37638 MW; EA17DC0CBDB8BA1 CRC64;

Query Match 31.7%; Score 57; DB 5; Length 345;
 Best Local Similarity 38.7%; Pred. No. 8.2;
 Matches 12; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

QY 5 ESSSTPSTGSGKTPSEDDRSQSRHMG 35
 DB 123 KESTNCGSFAPVQPTDKVDQSSGHMDTS 153

RESULT 10

Q9M010 PRELIMINARY; PRT; 621 AA.

AC Q9M010; 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Hypothetical 68.6 kDa protein.
 GN F7A7.180.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eustosids II; Brassicales; Brassicaceae; Arabidopsis.
 OK NCBI_TaxID=3702;

Q9M010 SEQUENCE FROM N.A.
 RA Bevan M., Pohl T., Weizenegger T., Bancroft I., Mewes H.W., Rudd S.,
 RA Lemcke K., Mayer K.F.X.;
 RN Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 RP [2]
 RA SEQUENCE FROM N.A.
 RP EU Arabidopsis sequencing project;
 RT Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AL161946; CAB82282.1;
 DR InterPro; IPR006552; Kelch_rep.
 DR Pfam; PF01344; Kelch; 6.
 DR SMART; SM00612; Kelch; 6.
 KW Hypothetical protein.
 SQ SEQUENCE 621 AA; 68579 MW; 259E20AFA74CE73P CRC64;

Query Match 31.1%; Score 56; DB 10; Length 621;
 Best Local Similarity 38.7%; Pred. No. 22;
 Matches 12; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

Oy 1 SFOSESTPTSGFGSKETPSDDRSQSRHMG 31
 Db 475 SVEHKSIVAVGVGDKETLNTAERFDPREH 505

RESULT 11

Oy 097T67 PRELIMINARY; PRT; 89 AA.

AC 097T67;
 DT 01-OCT-2001 (TREMBlrel. 18, Created)
 DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Hypothetical protein SP0088.
 GN SP0088.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TIGR4;
 RX MEDLINE=21357209; PubMed=11463916;

RA Tetteijn H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
 RA Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
 RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
 RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
 RA Holtzapple E., Khouri H., Wolf A.M., Utecherback T.R., Hansen C.L.,
 RA McDonald L.A., Feldguy T.V., Angluoli S., Dickinson T., Hickey E.K.,
 RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
 RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Frazer C.M.;
 RA "Complete genome sequence of a virulent isolate of Streptococcus
 RT pneumoniae.";
 RT Science 293:498-506(2001).
 DR EMBL; AE007326; AAK74275.1; -.
 DR TIGR; SP0088; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 89 AA; 9326 MW; 3D2136B4CC51426 CRC64;

Query Match 30.6%; Score 55; DB 16; Length 89;
 Best Local Similarity 31.2%; Pred. No. 3.5;
 Matches 10; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

Oy 3 QSESSTPTSGFGSKETPSDDRSQSRHMG 34
 Db 55 ESQATKDGQGTAGKETAGDESNATQOEISQ 86

RESULT 12

Oy 08CZC2 PRELIMINARY; PRT; 89 AA.

AC 08CZC2;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Hypothetical protein.
 GN SP0080.
 OS Streptococcus pneumoniae (strain ATCC BAA-295 / R6).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=171101;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21429245; PubMed=11544234;
 RA Hoskins J., Alborn W.E. Jr., Arnold J., Blaszczak L.C., Burgett S.,
 RA Dehoff B.S., Battem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
 RA Gilmour R., Glaes J.S., Khoja H., Kraft A.R., Lagace R.E.,
 RA LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
 RA McAhren S.M., McHenry M., Mcleaster K., Mundy C.W., Niclas T.I.,

RA Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rockey P.,
 RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
 RA Zook C.A., Baltz R.H., Jaekunas S.R., Rosteck P.R. Jr., Skatrud P.L.,
 RA Glas J.I.;
 RA "Genome of the bacterium Streptococcus pneumoniae strain R6.";
 RT J. Bacteriol. 183:5709-5717(2001).
 RL EMBL; AE008392; AAK98884.1; -.
 DR KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 89 AA; 9326 MW; 3D2136B4CC51426 CRC64;

Query Match 30.6%; Score 55; DB 16; Length 89;
 Best Local Similarity 31.2%; Pred. No. 3.5;
 Matches 10; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

Oy 3 QSESSTPTSGFGSKETPSDDRSQSRHMG 34
 Db 55 ESQATKDGQGTAGKETAGDESNATQOEISQ 86

RESULT 13

Oy 081WH6 PRELIMINARY; PRT; 287 AA.

AC 081WH6;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Hypothetical protein (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Bladder;
 RA Strauberg R.;
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC039706; AAH39706.1; -.
 KW Hypothetical protein.
 FT NON_TER 287
 SQ SEQUENCE 287 AA; 33143 MW; 485FECAB1F191948 CRC64;

Query Match 30.6%; Score 55; DB 4; Length 287;
 Best Local Similarity 42.9%; Pred. No. 13;
 Matches 12; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

Oy 3 QSESSTPTSGFGSKETPSDDRSQSRH 30
 Db 247 ESENEITSVGASGDGDSDDDEQEE 274

RESULT 14

Oy 096J32 PRELIMINARY; PRT; 656 AA.

AC 096J32;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Voltage-dependent calcium channel alpha13 subunit (Fragment).
 GN CA-ALPHA-1T OR CG4222 OR CG15899.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Miltman S.;
 RT "Exon organization of Ca2+ channel alpha13 subunit genes.";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF31281; AAK54737.1; -.
 DR FLYBase; FBgn0029846; Ca-alpha-1T.
 FT NON_TER 1
 FT NON_TER 656

SEQ SEQUENCE 656 AA; 72626 MW; CE139DB829EA1CFF CRC64;

Query Match 30.6%; Score 55; DB 5; Length 656;

Best Local Similarity 37.0%; Pred. No. 32;

Matches 10; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

Qy 4 SSSSTPSTGFGFGKTPSEDDRSQSRE 30
 Db 100 ASSAAPASAGTSGASAFGERDRDRD 126

RESULT 15

Q8IVX7 PRELIMINARY; PRT; 536 AA.
 ID Q8IVX7;
 AC Q8IVX7;
 DT 01-MAR-2003 (TRENBLREL. 23, Created)
 DT 01-MAR-2003 (TRENBLREL. 23, Last sequence update)
 DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)
 DE Similar to eyes absent homolog 3 (Drosophila).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC041667; AAH41667.1; -
 SQ SEQUENCE 536 AA; 58697 MW; 7F502740058138F7 CRC64;

Query Match 30.3%; Score 54.5; DB 4; Length 536;

Best Local Similarity 31.9%; Pred. No. 30;

Matches 15; Conservative 6; Mismatches 11; Indels 15; Gaps 1;

Qy 1 SFOSE-----SSTPSTGFGFGKTPSEDDRSQSREHM 32
 Db 190 TYQSEKPSVMAAPAPAAQRLSSGDPSTPSLSQTPSKDTDDQSRKMM 236

Search completed: July 24, 2003, 11:57:03
 Job time : 11.7892 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 11:58:19 ; Search time 10.4451 Seconds
(without alignments)
623.047 Million cell updates/sec

Title: US-09-991-681-28

Perfect score: 41
Sequence: 1 EDERSTYSSQCCSSEDEDI.....ETNAQVSPRGKRRQWRAR 41

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1107863 seqs, 158726573 residues

Word size: 0

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

A_Geneseq_19Jun03:*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	41	23	AAO19166
2	41	100.0	41	23	AAO19166
3	41	100.0	518	19	AAW85472
4	41	100.0	518	23	AAO19165
5	41	100.0	518	23	AAW85472
6	41	100.0	1807	22	ABG09728
7	41	100.0	1839	23	ABP64835
8	41	100.0	1982	22	ABG09731
9	13	31.7	192	20	AAV48248

10	7	17.1	111	22	AAO13359	Human polypeptide
11	7	17.1	150	20	AAV48216	Human prostate can
12	7	17.1	512	23	AAE19559	Heres simplex vir
13	7	17.1	610	22	AB93568	Human protein sequ
14	7	17.1	790	21	AAO1923	Arabidopsis thalia
15	7	17.1	794	21	AAO1922	Arabidopsis thalia
16	7	17.1	814	21	AAO1921	Arabidopsis thalia
17	7	17.1	898	19	AAW78898	Rat UNC-5 homology
18	7	17.1	898	23	AAU97900	Rat netrin binding
19	7	17.1	898	23	AAU10543	Rat netrin receptor
20	7	17.1	1404	23	AAE20581	Proliferation pote
21	6	14.6	9	14	AAE27416	PK40 TAU/neurofila
22	6	14.6	9	20	AAE27416	PK40 protein kinase
23	6	14.6	9	23	AAE51124	Bovine TAU/neurofi
24	6	14.6	14	23	AAE24500	Human RATT16 case
25	6	14.6	15	17	AAE24501	Human RATT16 case
26	6	14.6	15	17	AAE24502	Human RATT16 case
27	6	14.6	16	24	AAE24503	Peptide #6, anti-M
28	6	14.6	16	24	AAE24504	G protein-coupled
29	6	14.6	16	24	AAE24505	Mnc1 acidic domain
30	6	14.6	21	23	AAU97857	Protein inhibitor
31	6	14.6	35	23	ABH79717	Extracellular sign
32	6	14.6	61	22	AAU41590	Protonibacterium
33	6	14.6	61	22	AAU66177	Protonibacterium
34	6	14.6	67	20	AAV12457	Human 5' EST seque
35	6	14.6	67	23	ABP06485	Human ORFX protein
36	6	14.6	71	24	ABR41286	Human DTHP intrac
37	6	14.6	80	23	ABR66596	Omega-conopeptide
38	6	14.6	81	22	AAU51203	Protonibacterium
39	6	14.6	82	22	AAO12252	Human polypeptide
40	6	14.6	85	22	AAU64469	Protonibacterium
41	6	14.6	93	22	AAO10196	Human polypeptide
42	6	14.6	97	22	ABG06996	Novel human diagno
43	6	14.6	100	22	ABG12857	Novel human diagno
44	6	14.6	100	22	ABG12858	Novel human diagno
45	6	14.6	105	22	AAO03991	Human polypeptide

ALIGNMENTS

RESULT 1	AAO19166	standard; Protein: 41 AA.
ID	AAO19166	
XX	XX	
AC	AAO19166	
XX	XX	
DT	27-NOV-2002	(first entry)
XX	XX	
DE	Human prostate-specific PS18 protein fragment #2.	
XX	XX	
KW	Human, prostate; prostate-specific sequence; prostate cancer; PS18;	
KW	EST; expressed sequence tag; cytosolic; gene therapy.	
XX	XX	
OS	Homo sapiens.	
XX	XX	
PN	US2002086316-A1.	
XX	XX	
PD	04-JUL-2002.	
XX	XX	
PF	26-NOV-2001; 2001US-0991681.	
XX	XX	
PR	23-APR-1998; 98US-0065383.	
XX	XX	
PR	23-APR-1997; 97US-0842385.	
XX	XX	
PA	(BIL/) BILINGEL P A.	
PA	(COHE/) COHEN M.	
PA	(COLP/) COLPITTS T L.	
PA	(FRIE/) FRIEDMAN P N.	
PA	(GORD/) GORDAN J E N.	
PA	(GRAN/) GRANADOS E N.	
PA	(HODG/) HODGES S C.	
PA	(KLAS/) KLAS M R.	

```

PA (KRAT/) KRATOCHVIL J D.
PA (ROBE/) ROBERTS-RAPP L.
PA (RUSS/) RUSSELL J C.
PA (STRO/) STROUPE S D.
XX
PI Billingel PA, Cohen M, Colpits TL, Friedman PN, Gordon J;
PI Granados EN, Hodges SC, Klass MR, Kratochvil JD, Roberts-Rapp L;
PI Russell JC, Stroupe SD;
XX
DR WPI; 2002-665429/71.
XX
PT Novel PS118 polypeptide for detecting, diagnosing, staging, monitoring,
PT prognosticating, preventing, treating, or determining predisposition of
PT individual to diseases and conditions of prostate, e.g. prostate
PT cancer -
XX
PS Claim 17; Page 43-44; 58pp; English.
XX
CC The present invention relates to a number of prostate-specific sequences
CC derived from the human PS118 gene. These can be used in the detection,
CC monitoring and treatment of prostate diseases, particularly prostate
CC cancer. The present sequence is a PS118 protein fragment of the
CC invention. The coding sequences of the invention were isolated from a
CC prostate tissue expressed sequence tag (EST) library.
XX
SQ Sequence 41 AA;
XX
Query Match 100.0%; Score 41; DB 23; Length 41;
Best Local Similarity 100.0%; Pred. No. 7.3e-36;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
OY 1 EDDERSTDSSQCCSEDEDFEETAOVSPPRGKERKROWAR 41
DB 1 EDDERSTDSSQCCSEDEDFEETAOVSPPRGKERKROWAR 41
XX
RESULT 2
ID AAM50810 standard; Protein; 41 AA.
XX
AC AAM50810;
XX
DT 01-MAY-2002 (first entry)
XX
DE PS118 prostate marker immunogenic polypeptide.
XX
KW PS118; prostate; marker; prostate cancer; tumour; metastasis;
KW benign prostatic hyperplasia; prostatic intraepithelial neoplasia;
KW prostaticitis; human; diagnosis; therapy; vaccine; immunogen.
XX
OS Homo sapiens.
XX
FN US2001055758-A1.
XX
PD 27-DEC-2001.
XX
PF 23-APR-1998; 98US-0065383.
XX
PR 23-APR-1997; 97US-0842385.
XX
PA (BILL/) BILLING-MEDEL P A.
PA (COHE/) COHEN M.
PA (COPL/) COPLPITTS T L.
PA (FRIE/) FRIEDMAN P N.
PA (GORD/) GORDON J.
PA (GRAN/) GRANADOS E N.
PA (HODG/) HODGES S C.
PA (KLAS/) KLAS M R.
PA (KRAT/) KRATOCHVIL J D.
PA (ROBE/) ROBERTS-RAPP L.
XX
PI Billing-Medel PA, Cohen M, Colpits TL, Friedman PN, Gordon J;
PI Granados EN, Hodges SC, Klass MR, Kratochvil JD, Roberts-Rapp L;

```

```

PI Russell JC, Stroupe SD;
XX
DR WPI; 2002-187683/24.
XX
PT Detecting presence of target PS118 polynucleotide in test sample,
PT useful for detecting, diagnosing, staging, monitoring, prognosticating,
PT preventing or treating or determining predisposition to prostate
PT disease -
XX
PS Claim 17; Page 43; 57pp; English.
XX
CC The present sequence is that of an immunogenic polypeptide
CC comprising amino acids 184-224 of human prostate-specific PS118
CC polypeptide (see AAM50809). A PS118 consensus sequence (see
CC ABA91851) is found at least 12 times more often in prostate than
CC in non-prostate tissue. PS118 polypeptides, polynucleotides,
CC antibodies, agonists and inhibitors are useful for detecting,
CC diagnosing, staging, monitoring, prognosticating, preventing and
CC treating (including by genetic immunisation), or determining the
CC predisposition of an individual to, diseases and conditions of the
CC prostate, such as benign prostatic hyperplasia, prostaticitis,
CC prostatic intraepithelial neoplasia, prostate cancer, tumours and
CC metastases. The PS118 polypeptides can be produced by expression
CC of PS118 polynucleotides in transfected host cells, and
CC immunogenic peptides are useful for raising PS118-specific
CC antibodies of diagnostic use. The methods and reagents of the
CC invention may provide an early means of detecting diseases of the
CC prostate and may also provide new markers which can differentiate
CC between the clinically important and unimportant prostate cancers
CC without the use of surgery.
XX
SQ Sequence 41 AA;
XX
Query Match 100.0%; Score 41; DB 23; Length 41;
Best Local Similarity 100.0%; Pred. No. 7.3e-36;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
OY 1 EDDERSTDSSQCCSEDEDFEETAOVSPPRGKERKROWAR 41
DB 1 EDDERSTDSSQCCSEDEDFEETAOVSPPRGKERKROWAR 41
XX
RESULT 3
ID AAM85472 standard; Protein; 518 AA.
XX
AC AAM85472;
XX
DT 25-MAR-2003 (updated)
DT 25-FEB-1999 (first entry)
XX
DE PS118 protein encoded by consensus sequence.
XX
KW EST clone; PS118; prostate tumour tissue; prostatic disease; cancer.
XX
OS Homo sapiens.
XX
FN WO9848054-A1.
XX
PD 29-OCT-1998.
XX
PF 23-APR-1998; 98WO-US08239.
XX
PR 23-APR-1997; 97US-0842385.
XX
PA (ABBO ) ABBOTT LAB.
XX
PI Billing-medel PA, Cohen M, Colpits TL, Friedman PN, Gordon J;
PI Granados EN, Hodges SC, Klass MR, Kratochvil JD, Robertterapp L;
PI Russell JC, Stroupe SD;
XX
DR WPI; 1998-610000/51.
DR N-PSDB; AAV82812.

```


PS Claim 17; Page 42-43; 57pp; English.

XX The present sequence is that of a human prostate-specific PS118
CC polypeptide, as predicted from a partial consensus cDNA sequence
CC (see ABA91651), and lacking the N-terminal region. The PS118
CC consensus sequence is found at least 12 times more often in
CC prostate than in non-prostate tissue. PS118 polypeptides,
CC including derivatives of the present sequence, polynucleotides,
CC antibodies, agonists and inhibitors are useful for detecting,
CC diagnosing, staging, monitoring, prognosticating, preventing and
CC treating (including by genetic immunisation), or determining the
CC predisposition of an individual to, diseases and conditions of the
CC prostate, such as benign prostatic hyperplasia, prostatitis,
CC prostatic intraepithelial neoplasia, prostate cancer, tumours and
CC metastases. The PS118 polypeptides can be produced by expression
CC of PS118 polynucleotides in transfected host cells. The methods
CC and reagents of the invention may provide an early means of
CC detecting diseases of the prostate and may also provide new markers
CC which can differentiate between the clinically important and
CC unimportant prostate cancers without the use of surgery.

XX Sequence 518 AA;

SO Query Match 100.0%; Score 41; DB 23; Length 518;
Best Local Similarity 100.0%; Pred. No. 6.1e-35;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDERSTDSQQCSSEDEDFEETAAVSPPRGKKRQWRAR 41
Db 184 EDDERSTDSQQCSSEDEDFEETAAVSPPRGKKRQWRAR 224

RESULT 6
ABG09728
ID ABG09728 standard; Protein; 1807 AA.

XX ABG09728;
XX
XX 13-FEB-2002 (first entry)
XX
XX Novel human diagnostic protein #9719.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX PN
XX PD 11-OCT-2001.
XX
XX PF 30-MAR-2001; 2001WO-US08631.
XX
XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX
XX PA (HYSE-) HYSEQ INC.
XX
XX Drmanac RT; Liu C, Tang YT;
XX
XX PI WPI; 2001-639362/73.
XX DR N-PSDB; AAS73915.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity -
XX
XX PS Claim 20; SEQ ID No 40087; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polynucleotide chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1807 AA;

SO Query Match 100.0%; Score 41; DB 22; Length 1807;
Best Local Similarity 100.0%; Pred. No. 1.8e-34;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDERSTDSQQCSSEDEDFEETAAVSPPRGKKRQWRAR 41
Db 1473 EDDERSTDSQQCSSEDEDFEETAAVSPPRGKKRQWRAR 1513

RESULT 7
ABP64835
ID ABP64835 standard; Protein; 1839 AA.

XX ABP64835;
XX
XX 25-FEB-2003 (first entry)
XX
XX Human protein SEQ ID 495.
XX
XX Human; expressed sequence tag; EST;
XX haematopoietic disorder; central nervous system disease; viral infection;
XX peripheral nervous system disease; non-healing wound; infectious disease;
XX immune deficiency; immune disorder; bacterial infection; allergy; cancer;
XX fungal infection; autoimmune disorder; coagulation disorder; noctropic;
XX antiallergic; antiinflammatory; immunosuppressive; neuroprotective;
XX cytostatic; haemostatic; virocidic; antibacterial; fungicide;
XX immunostimulant; cerebroprotective.

XX OS Homo sapiens.
XX
XX PN WO200259260-A2.
XX
XX PD 01-AUG-2002.
XX
XX PF 16-NOV-2001; 2001WO-US42950.
XX
XX PR 17-NOV-2000; 2000US-0714936.
XX
XX PA (HYSE-) HYSEQ INC.
XX
XX Tang YT, Goodrich RW, Liu C, Zhou P, Aundri V, Zhang J, Zhao QA,
XX PI Ren F, Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX
XX WPI; 2002-590824/63.
XX DR N-PSDB; ABQ99421.
XX
XX New isolated polynucleotide, useful in research, diagnostic or
XX PT therapeutic methods, e.g. preventing or treating disorders involving
XX aberrant protein expression or biological activity -
XX
XX PS Claim 20; SEQ ID 495; 394pp; English.

The present invention relates to novel human coding sequences (AB093268-AB093608) and proteins (ABP64682-ABP65022). The sequences are useful in therapeutic, diagnostic and research methods. The polynucleotides may be used in the field of molecular biology as hybridisation probes, primers for PCR, for chromosome and gene mapping, for the recombinant production of protein, or in generation of anti-sense DNA or RNA. The polynucleotides are useful in diagnostics as expressed sequence tags (ESTs) for identifying expressed genes or for physical mapping of the human genome. The proteins may be used as molecular weight markers, or as nutritional sources or supplements. The proteins may be used to maintain and expand cell population in a totipotent or pluripotent state useful for re-engineering damaged or diseased tissues, transplantation, manufacture of bio-pharmaceuticals or the development of bio-sensors. The polynucleotides and proteins are useful for preventing, treating or ameliorating disorders involving aberrant protein expression or biological activity, e.g. haematopoietic disorders, central/peripheral nervous system diseases, mechanical and traumatic disorders, non-healing wounds, immune deficiencies and disorders, infectious diseases caused by viral, bacterial or fungal infection, autoimmune disorders, allergic reactions and conditions, coagulation disorders, or cancer. The polynucleotide sequences of the invention were assembled from ESTs isolated mainly by sequencing by hybridisation, and in some cases, sequences obtained from one or more public databases. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 1839 AA;

Query Match 100.0%; Score 41; DB 23; Length 1839;
Best Local Similarity 100.0%; Pred. No. 1.8e-34;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EDDERSTDSQCCSSEDEDFEETAAVSPPRGKEKROWRAR 41
Db 1505 EDDERSTDSQCCSSEDEDFEETAAVSPPRGKEKROWRAR 1545

RESULT 8

ABG09731 ID ABG09731 standard; Protein; 1982 AA.

AC ABG09731;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #9722.

KM Human; chromosome mapping; gene mapping; gene therapy; forensic;

KM food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

PI Dzmanac RT, Liu C, Tang YT;

DR WPI; 2001-633962/73.

DR N-PSDB; AAS73918.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity

XX Claim 20; SEQ ID No 40090; 103pp; English.

PS The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polynucleotide chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

Sequence 1982 AA;

Query Match 100.0%; Score 41; DB 22; Length 1982;
Best Local Similarity 100.0%; Pred. No. 1.9e-34;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EDDERSTDSQCCSSEDEDFEETAAVSPPRGKEKROWRAR 41
Db 1587 EDDERSTDSQCCSSEDEDFEETAAVSPPRGKEKROWRAR 1627

RESULT 9

AAV48248 ID AAV48248 standard; Protein; 192 AA.

AC AAV48248;

DT 08-DEC-1999 (first entry)

DE Human prostate cancer-associated protein 34.

KM Expressed sequence tag; EST; prostate tumor; antitumor; treatment;

KM gene therapy; tissue specificity human.

OS Homo sapiens.

PN DE1981193-A1.

PD 16-SEP-1999.

PF 10-MAR-1998; 98DE-1011193.

PR 10-MAR-1998; 98DE-1011193.

XX (META-) METAGEN GES GENOMFORSCHUNG MBH.

PI Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosenthal A;

DR WPI; 1999-519628/44.

DR N-PSDB; AAZ33451.

PT New nucleic acid expressed at high level in prostatic tumor tissue and
PT encoded polypeptides, useful for treating cancer and screening for
PT therapeutic agents

PS Claim 22; 128; 166pp; German.

XX This invention describes novel nucleic acid sequences (A) that are

CC expressed at high level in prostatic tumor tissue and encode gene
 CC products of their fragments. The products of the invention have
 CC antitumor activity. Polypeptides (I) encoded by (A) are used: (i) for
 CC identifying agents for treatment of prostatic cancer and (ii) for
 CC therapy of prostate cancer, optionally where expressed by gene therapy
 CC methods. (A) is also used to isolate full-length genes (for gene therapy)
 CC and for recombinant production of (i), which can be used to raise
 CC specific antibodies. (A) are identified by assembly of ESTs (expressed
 CC sequence tags) before they are analyzed for expression pattern (tissue
 CC specificity). This approach eliminates many of the false results, as
 CC regards tissue specificity, associated with known methods that use
 CC single (usually short) ESTs. AAY48215-Y48303 represent protein fragments
 CC encoded by the expressed sequence tags described in the method of the
 CC invention.

SQ Sequence 192 AA;
 Query Match 31.7%; Score 13; DB 20; Length 192;
 Best Local Similarity 100.0%; Pred. No. 8e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDERSTDSQQC 13
 |||||
 DB 180 EDDERSTDSQQC 192

RESULT 10
 AA011359
 ID AA011359 standard; Protein; 111 AA.
 AC AA011359;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 25251.
 XX
 XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation.
 XX
 OS Homo sapiens.
 XX
 PN WO200164835-A2.
 XX
 PD 07-SEP-2001.
 XX
 PF 26-FEB-2001; 2001WO-US04927.
 XX
 PR 28-FEB-2000; 2000US-0515126.
 PR 18-MAY-2000; 2000US-0577409.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YF, Liu C, Drmanac RT;
 XX
 DR MPI: 2001-514838/56.
 DR N-PSDB; AA191290.
 XX
 PT Isolated nucleic acids and polypeptides, useful for preventing
 PT diagnosing and treating e.g. leukaemia, inflammation and immune
 PT disorders -
 XX
 PS Claim 20; SEQ ID NO 25251; 1399bp + Sequence Listing; English.
 CC The invention relates to human polynucleotides (AA179941-AA193841) and
 CC the encoded proteins (AA000010-AA013910) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and

CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 111 AA;
 Query Match 17.1%; Score 7; DB 22; Length 111;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 VSPPRGK 33
 |||||
 DB 28 VSPPRGK 34

RESULT 11
 AA48216
 ID AA48216 standard; Protein; 180 AA.
 AC AA48216;
 XX
 DT 08-DEC-1999 (first entry)
 XX
 DE Human prostate cancer-associated protein 2.
 KW Expressed sequence tag; EST; prostate tumor; antitumor; treatment;
 KW gene therapy; tissue specificity human.
 XX
 OS Homo sapiens.
 XX
 PN DE1981193-A1.
 XX
 PD 16-SEP-1999.
 XX
 PF 10-MAR-1998; 98DE-1011193.
 PR 10-MAR-1998; 98DE-1011193.
 XX
 PA (META-) METAGEN GES GENOMFORSCHUNG MBH.
 XX
 PI Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosenthal A;
 XX
 DR MPI: 1999-519628/44.
 DR N-PSDB; AA233424.
 XX
 PT New nucleic acid expressed at high level in prostatic tumor tissue and
 PT encoded polypeptides, useful for treating cancer and screening for
 PT therapeutic agents -
 XX
 PS Claim 22; 112-113; 166bp; German.
 XX
 CC This invention describes novel nucleic acid sequences (A) that are
 CC expressed at high level in prostatic tumor tissue and encode gene
 CC products of their fragments. The products of the invention have
 CC antitumor activity. Polypeptides (I) encoded by (A) are used: (i) for
 CC identifying agents for treatment of prostatic cancer and (ii) for
 CC therapy of prostate cancer, optionally where expressed by gene therapy
 CC methods. (A) is also used to isolate full-length genes (for gene therapy)
 CC and for recombinant production of (i), which can be used to raise
 CC specific antibodies. (A) are identified by assembly of ESTs (expressed
 CC sequence tags) before they are analyzed for expression pattern (tissue
 CC specificity). This approach eliminates many of the false results, as
 CC regards tissue specificity, associated with known methods that use
 CC single (usually short) ESTs. AAY48215-Y48303 represent protein fragments
 CC encoded by the expressed sequence tags described in the method of the
 CC invention.

SQ Sequence 180 AA;
 Query Match 17.1%; Score 7; DB 20; Length 180;

Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDDERST 7
DB 174 EDDERST 180

RESULT 12

AAE19559 standard; protein; 512 AA.

XX AAE19559;

XX 31-MAY-2002 (first entry)

XX Herpes simplex virus type 2 (HSV-2) IE63 protein homologue.

XX Herpes virus infection; detection; therapy; IE63 protein; antiviral.

XX Herpes simplex virus type 2.

XX W0200204492-A2.

XX 17-JAN-2002.

XX 11-JUL-2001; 2001WO-GB03114.

XX 11-JUL-2000; 2000GB-0016890.

XX (UNIU) UNIV GLASGOW.

XX Clements JB, Maclean AR;

XX WPI; 2002-226983/28.

XX Detecting an agent useful for treating herpes virus infection comprises determining any change in a polypeptide/zinc complex in the presence of the test agent -

XX Claim 6; Fig 1; 43pp; English.

XX The invention relates to a method of detecting an agent for use in the treatment of herpes virus infection. The method comprises forming a herpes virus polypeptide/zinc complex; adding a test agent to the complex; and detecting any change in the complex.
XX The invention also relates to the use of known agents, such as 2',5'-dithiobisbenzamide (DTBA) and azodicarbonamide (ADA), and unknown agents for the manufacture of a medicament for the treatment of herpes virus infection. The method is useful for detecting agents for use in the treatment of herpes virus infection. The present sequence is Herpes simplex virus type 2 (HSV2) IE63 protein homologue.
XX HSV2 belongs to alphaherpes virus sub-family. IE63 protein is also called ICP27.

XX Sequence 512 AA;

Query Match 17.1%; Score 7; DB 23; Length 512;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 CSSEDED 19
DB 44 CSSEDED 50

RESULT 13

AAAB93588 standard; Protein; 610 AA.

XX AAAB93588;

XX 26-JUN-2001 (first entry)

XX Human protein sequence SEQ ID NO:13012.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX Homo sapiens.

XX EPI074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99JP-0300253.

XX 11-JAN-2000; 2000JP-0118776.

XX 02-MAY-2000; 2000JP-0183767.

XX 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Oca T, Isogai T, Nishikawa T, Hayashi K, Saico K, Yamamoto J;

XX Ienit S, Sugiyama T, Wakamatsu A, Negai K, Otsuki T;

XX WPI; 2001-318749/34.

XX Claim 8; SEQ ID 13012; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any special methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AA892446 to AA895893 represent human amino acid sequences; and AAH1629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

XX Sequence 610 AA;

Query Match 17.1%; Score 7; DB 22; Length 610;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 14 SSEDEDI 20
DB 194 SSEDEDI 200

RESULT 14

AAAG41923 standard; Protein; 790 AA.

XX AAAG41923;

XX AAAG41923;

DT 18-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 52219.
DE
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138647.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145224.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146386.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149358.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.

PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154778.
PR 22-SEP-1999; 99US-0155138.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0156559.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158233.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159299.
PR 13-OCT-1999; 99US-0159299.
PR 13-OCT-1999; 99US-0159299.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160777.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query March 17.1%; Score 7; DB 21; Length 790;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 27 VSPRGRK 33
Db 639 VSPRGRK 645

RESULT 15
AAG41922
ID AAG41922 standard; Protein; 794 AA.

XX AC AAG41922;

XX DT 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 52218.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hydridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX OS Arabidopsis thaliana.
XX PN EPI033405-A2.
XX FD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138647.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.

PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.

PR 13-SEP-1999; 99US-0153758.
 PR 13-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 17.1%; Score 7; DB 21; Length 794;
 Best Local Similarity 100.0%; Pred. No. 55;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 27 VSPPRGX 33
 DB 643 VSPPRGX 649

Search completed: July 24, 2003, 12:21:15
 Job time : 12.4451 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 12:18:29 ; Search time 4.74231 Seconds
(without alignments)
365.801 Million cell updates/sec

Title: US-09-991-681-28

Perfect score: 41
Sequence: 1 EDDERSTDSQCCSEDEDI.....EETAQVSPPRKEXRQWRAR 41

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310858 residues

Word size : 0

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents_AA.*

1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	41	US-09-065-383-28	Sequence 28, Appl
2	41	100.0	518	US-09-065-383-27	Sequence 27, Appl
3	7	17.1	898	US-08-808-982-5	Sequence 5, Appl
4	7	17.1	898	US-09-306-902A-5	Sequence 5, Appl
5	7	17.1	1404	US-08-801-308-1	Sequence 1, Appl
6	6	14.6	9	US-08-480-793-9	Sequence 9, Appl
7	6	14.6	9	PCT-US92-05825A-9	Sequence 9, Appl
8	6	14.6	14	US-08-622-277A-16	Sequence 16, Appl
9	6	14.6	15	US-07-893-929A-7	Sequence 5, Appl
10	6	14.6	130	PCT-US92-10344-7	Sequence 7, Appl
11	6	14.6	130	US-07-893-929A-1	Sequence 1, Appl
12	6	14.6	131	PCT-US92-10344-1	Sequence 1, Appl
13	6	14.6	131	US-07-893-929A-5	Sequence 5, Appl
14	6	14.6	132	US-07-893-929A-5	Sequence 5, Appl
15	6	14.6	134	US-07-893-929A-2	Sequence 2, Appl
16	6	14.6	134	PCT-US92-10344-2	Sequence 2, Appl
17	6	14.6	159	US-09-252-991A-31988	Sequence 31988, A
18	6	14.6	282	US-08-972-902-7	Sequence 7, Appl
19	6	14.6	282	US-09-520-207-7	Sequence 7, Appl
20	6	14.6	323	US-09-252-991A-21180	Sequence 21180, A
21	6	14.6	358	US-08-176-620A-12	Sequence 12, Appl
22	6	14.6	358	US-08-461-985-12	Sequence 12, Appl
23	6	14.6	358	US-09-457-040B-4	Sequence 4, Appl
24	6	14.6	360	US-08-622-277A-14	Sequence 14, Appl
25	6	14.6	360	US-09-025-580-2	Sequence 2, Appl
26	6	14.6	360	US-09-457-040B-3	Sequence 3, Appl
27	6	14.6	360	US-09-457-040B-3	Sequence 3, Appl

28	6	14.6	364	US-09-457-040B-34	Sequence 34, Appl
29	6	14.6	365	US-09-457-040B-35	Sequence 35, Appl
30	6	14.6	371	US-09-252-991A-29793	Sequence 29793, A
31	6	14.6	380	US-08-459-953A-9	Sequence 9, Appl
32	6	14.6	380	US-09-393-212-9	Sequence 9, Appl
33	6	14.6	415	US-08-176-620A-4	Sequence 4, Appl
34	6	14.6	415	US-08-463-862-4	Sequence 4, Appl
35	6	14.6	415	US-08-463-862-4	Sequence 4, Appl
36	6	14.6	415	US-08-458-887-4	Sequence 4, Appl
37	6	14.6	415	US-08-932-787B-4	Sequence 4, Appl
38	6	14.6	415	US-08-932-012C-4	Sequence 4, Appl
39	6	14.6	415	US-08-888-818C-4	Sequence 4, Appl
40	6	14.6	502	US-08-484-840-3	Sequence 3, Appl
41	6	14.6	502	US-08-483-094-3	Sequence 3, Appl
42	6	14.6	511	US-09-252-991A-19753	Sequence 19753, A
43	6	14.6	543	US-09-328-352-5845	Sequence 5845, Ap
44	6	14.6	567	US-09-205-258-573	Sequence 573, App
45	6	14.6	569	US-09-137-223A-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-09-065-383-28
Sequence 28, Application US/09065383
Patent No. 6391543
GENERAL INFORMATION:
APPLICANT: BILTING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GORDON, JULIAN
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KLAS, MICHAEL R.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: ROBERTS-RAPP, LISA
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESSES:
ADDRESSER: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Fastseq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/065,383
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/842,385
FILING DATE: 23-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6084, US, P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 41 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: NO. 6391543e
US-09-065-383-28

Query Match 100.0%; Score 41; DB 4; Length 41;
Best Local Similarity 100.0%; Pred. No. 1,4e-35;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EDDERSTDSSQCCSSEDEDFEETAOVSPPRGKEXKQWRAR 41
Db 1 EDDERSTDSSQCCSSEDEDFEETAOVSPPRGKEXKQWRAR 41

RESULT 2
US-09-065-383-27
Sequence 27, Application US/09065383
Patent No. 6391543

GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GORDON, JULIAN
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KLAAS, MICHAEL R.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: ROBERTS-RAPP, LISA
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
IN THE TITRATION OF PROSTATE
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/065,383
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/842,385
FILING DATE: 23-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6084.US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 518 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: NO. 6391543e
US-09-065-383-27

Query Match 100.0%; Score 41; DB 4; Length 518;
Best Local Similarity 100.0%; Pred. No. 1.2e-34;

Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 EDDERSTDSSQCCSSEDEDFEETAOVSPPRGKEXKQWRAR 41
Db 184 EDDERSTDSSQCCSSEDEDFEETAOVSPPRGKEXKQWRAR 224

RESULT 3
US-08-808-982-5
Sequence 5, Application US/08808982
Patent No. 5939271

GENERAL INFORMATION:
APPLICANT: Tessier-Lavigne, Marc
APPLICANT: Leonardo, E. David
APPLICANT: Hink, Lindsey
APPLICANT: Masu, Masayuki
APPLICANT: Kazuko, Keino-Masu
TITLE OF INVENTION: Netrin Receptors
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/808,982
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UC96-217
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 898 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-808-982-5

Query Match 17.1%; Score 7; DB 2; Length 898;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 ERSTDSS 10
Db 90 ERSTDSS 96

RESULT 4
US-09-306-902A-5
Sequence 5, Application US/09306902A
Patent No. 6277585

GENERAL INFORMATION:
APPLICANT: Tessier-Lavigne, Marc
APPLICANT: Leonardo, E. David
APPLICANT: Hink, Lindsey
APPLICANT: Masu, Masayuki
APPLICANT: Kazuko, Keino-Masu
TITLE OF INVENTION: Netrin Receptors
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:

ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/306,902A
FILING DATE: 07-May-1999
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UC96-217
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 898 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-306-902A-5
Query Match 17.1%; Score 7; DB 3; Length 898;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 4 ERSTDSS 10
Db 90 ERSTDSS 96
RESULT 5
US-08-801-308-1
Sequence 1, Application US/08801308
Patent No. 6368790
GENERAL INFORMATION:
APPLICANT: Scott, Robert E.
TITLE OF INVENTION: CDNA ENCODING P2P PROTEINS AND USE OF
TITLE OF INVENTION: P2P CDNA-DERIVED ANTIBODIES AND ANTISENSE REAGENTS IN
TITLE OF INVENTION: DETERMINING THE PROLIFERATIVE POTENTIAL OF NORMAL,
TITLE OF INVENTION: ABNORMAL AND CANCER CELLS IN ANIMALS AND HUMANS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Weiser & Associates, P.C.
STREET: 230 S. Fifteenth Street, Suite 500
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19102
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/801,308
FILING DATE: 18-FEB-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Weiser, Gerard J.
REGISTRATION NUMBER: 19,763
REFERENCE/DOCKET NUMBER: 372.6435P
TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-875-8383
TELEFAX: 215-875-8394
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1404 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-801-308-1
Query Match 17.1%; Score 7; DB 4; Length 1404;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 28 SPRGKE 34
Db 1288 SPRGKE 1294
RESULT 6
US-08-480-793-9
Sequence 9, Application US/08480793
Patent No. 595444
GENERAL INFORMATION:
APPLICANT: Ingram, Vernon M., Roder, Hanno M.
TITLE OF INVENTION: No. 595444e1 Tau/Neurofilament Protein
TITLE OF INVENTION: Kinases
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Massachusetts Institute of Technology
STREET: 28 Carleton Street
CITY: Cambridge
STATE: Massachusetts
COUNTRY: United States of America
ZIP: 02142
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch
COMPUTER: IBM-compatible
OPERATING SYSTEM: MS-DOS Version 3.3
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,793
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/912,293
FILING DATE: July 10, 1992
APPLICATION NUMBER: 07/742,880
FILING DATE: 9-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: Gates, Edward R.
REGISTRATION NUMBER: 31,616
REFERENCE/DOCKET NUMBER: M0656/7008
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-480-793-9
Query Match 14.6%; Score 6; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 20 IFEETA 25
Db 3 IFEETA 8
RESULT 7
PCT-US92-05825A-9
Sequence 9, Application PC/TUS9205825A

GENERAL INFORMATION:
APPLICANT: Ingram, Vernon M., Roder, Hanno M.
TITLE OF INVENTION: Novel Tau/Neurofilament Protein
TITLE OF INVENTION: Kinases
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Massachusetts Institute of Technology
STREET: 28 Carleton Street
CITY: Cambridge
STATE: Massachusetts
COUNTRY: United States of America
ZIP: 02142
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch
COMPUTER: IBM-compatible
OPERATING SYSTEM: MS-DOS Version 3.3
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/05825A
FILING DATE: 19920710
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/742,880
FILING DATE: 9-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: Gates, Edward R.
REGISTRATION NUMBER: 31,616
REFERENCE/DOCKET NUMBER: M0656/7008
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US92-05825A-9

Query Match 14.6%; Score 6; DB 5; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 IFEETA 25
Db 3 IFEETA 8

RESULT 8
US-08-622-277A-16
Sequence 16, Application US/08622277A
Patent No. 6001580
GENERAL INFORMATION:
APPLICANT: Tani, Akiyoshi
APPLICANT: Ichimori, Yuzo
TITLE OF INVENTION: Method For Assaying MAP Kinase
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/622,277A
FILING DATE: 27-MAR-1996
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 320577-1995

FILING DATE: 08-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 305456-1995
FILING DATE: 24-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 070125-1995
FILING DATE: 28-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Tooney, Kimberlin M.
REGISTRATION NUMBER: 35,391
REFERENCE/DOCKET NUMBER: 04221.0039-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)408-4000
TELEFAX: (202)408-4400
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-622-277A-16

Query Match 14.6%; Score 6; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 IFEETA 25
Db 1 IFEETA 6

RESULT 9
US-08-622-277A-5
Sequence 5, Application US/08622277A
Patent No. 6001580
GENERAL INFORMATION:
APPLICANT: Tani, Akiyoshi
APPLICANT: Ichimori, Yuzo
TITLE OF INVENTION: Method For Assaying MAP Kinase
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/622,277A
FILING DATE: 27-MAR-1996
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 320577-1995
FILING DATE: 08-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 305456-1995
FILING DATE: 24-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 070125-1995
FILING DATE: 28-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Tooney, Kimberlin M.
REGISTRATION NUMBER: 35,391
REFERENCE/DOCKET NUMBER: 04221.0039-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)408-4000

TELEFAX: (202)408-4400
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-622-277A-5

Query Match 14.6%; Score 6; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 6.4;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 IFEETA 25
Db 2 IFEETA 7

RESULT 10
US-07-893-929A-7
; Sequence 7, Application US/07893929A
; Patent No. 5336667
; GENERAL INFORMATION:
; APPLICANT: Kirby, Edward P.
; APPLICANT: Peng, Man-ling
; TITLE OF INVENTION: Alboagregins: Platelet
; TITLE OF INVENTION: Agonists Which Bind To Platelet
; TITLE OF INVENTION: Membrane Glycoprotein Ib
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Temple University - Of The Common-
; ADDRESSEE: wealth System of Higher Education
; STREET: 406 University Services Building
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/893,929A
; FILING DATE: 19920605
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/803,630
; FILING DATE: December 3, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 6056-126 (CIP) 1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: No. 5336667e
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 130 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
US-07-893-929A-7

Query Match 14.6%; Score 6; DB 1; Length 130;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 QQCSE 16
Db 75 QQCSE 80

RESULT 11
PCT-US92-10344-7
; Sequence 7, Application PC/TUS9210344
; GENERAL INFORMATION:
; APPLICANT: Kirby, Edward P.
; APPLICANT: Peng, Man-ling
; TITLE OF INVENTION: Alboagregins: Platelet
; TITLE OF INVENTION: Agonists Which Bind To Platelet
; TITLE OF INVENTION: Membrane Glycoprotein Ib
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Temple University - Of The Common-
; ADDRESSEE: wealth System of Higher Education
; STREET: 406 University Services Building
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/10344
; FILING DATE: 19921201
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/803,630
; FILING DATE: December 3, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 6056-126 (CIP) 1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: None
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 130 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
PCT-US92-10344-7

Query Match 14.6%; Score 6; DB 5; Length 130;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 QQCSE 16
Db 75 QQCSE 80

RESULT 12
US-07-893-929A-1
; Sequence 1, Application US/07893929A
; Patent No. 5336667
; GENERAL INFORMATION:
; APPLICANT: Kirby, Edward P.
; APPLICANT: Peng, Man-ling
; TITLE OF INVENTION: Alboagregins: Platelet
; TITLE OF INVENTION: Agonists Which Bind To Platelet
; TITLE OF INVENTION: Membrane Glycoprotein Ib
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Temple University - Of The Common-
; ADDRESSEE: wealth System of Higher Education
; STREET: 406 University Services Building
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.

ZIP: 19122
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/893,929A
FILING DATE: 19920605
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/803,630
FILING DATE: December 3, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Monaco, Daniel A.
REGISTRATION NUMBER: 30,480
REFERENCE/DOCKET NUMBER: 6056-126 (CIP) 1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-8383
TELEFAX: (215) 568-5549
TELEX: No. 5336667e
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 131 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
US-07-893-929A-1

Query Match 14.6%; Score 6; DB 1; Length 131;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 QOCSSSE 16
Db 77 QOCSSSE 82

RESULT 13
PCT-US92-10344-1
Sequence 1, Application PC/TUS9210344
GENERAL INFORMATION:
APPLICANT: Kirby, Edward P.
APPLICANT: Peng, Man-ling
TITLE OF INVENTION: Alboesgregins: Platelet
TITLE OF INVENTION: Agonists which Bind To Platelet
TITLE OF INVENTION: Membrane Glycoprotein Ib
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Temple University - Of The Common-
ADDRESS: wealth System of Higher Education
STREET: 406 University Services Building
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: U.S.A.
ZIP: 19122
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/10344
FILING DATE: 19921201
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/803,630
FILING DATE: December 3, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Monaco, Daniel A.
REGISTRATION NUMBER: 30,480
REFERENCE/DOCKET NUMBER: 6056-126 (CIP) 1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-8383

TELEFAX: (215) 568-5549
TELEX: None
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 131 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
PCT-US92-10344-1

Query Match 14.6%; Score 6; DB 5; Length 131;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 QOCSSSE 16
Db 77 QOCSSSE 82

RESULT 14
US-07-893-929A-5
Sequence 5, Application US/07893929A
Patent No. 5336667
GENERAL INFORMATION:
APPLICANT: Kirby, Edward P.
APPLICANT: Peng, Man-ling
TITLE OF INVENTION: Alboesgregins: Platelet
TITLE OF INVENTION: Agonists which Bind To Platelet
TITLE OF INVENTION: Membrane Glycoprotein Ib
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Temple University - Of The Common-
ADDRESS: wealth System of Higher Education
STREET: 406 University Services Building
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: U.S.A.
ZIP: 19122
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/893,929A
FILING DATE: 19920605
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/803,630
FILING DATE: December 3, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Monaco, Daniel A.
REGISTRATION NUMBER: 30,480
REFERENCE/DOCKET NUMBER: 6056-126 (CIP) 1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-8383
TELEFAX: (215) 568-5549
TELEX: No. 5336667e
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 132 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
US-07-893-929A-5

Query Match 14.6%; Score 6; DB 1; Length 132;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 QOCSSSE 16
Db 77 QOCSSSE 82

RESULT 15
PCT-US92-10344-5
; Sequence 5, Application PC/TUS9210344
; GENERAL INFORMATION:
; APPLICANT: Kirby, Edward P.
; APPLICANT: Peng, Man-ling
; TITLE OF INVENTION: Alboaggregins: Platelet
; TITLE OF INVENTION: Agonists Which Bind To Platelet
; TITLE OF INVENTION: Membrane Glycoprotein Ib
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Temple University - Of The Common-
; ADDRESSEE: wealth System of Higher Education
; STREET: 406 University Services Building
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 KB
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/10344
; FILING DATE: 19921201
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/803,630
; FILING DATE: December 3, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 6056-126 (CIP) 1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: None
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 132 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; PCT-US92-10344-5

Query Match 14.6%; Score 6; DB 5; Length 132;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 QCCSSE 16
|||
Db 77 QCCSSE 82

Search completed: July 24, 2003, 12:28:37
Job time : 5.74231 secs

PAGE BLANK (verso)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 12:26:04 ; Search time 6.9634 Seconds
(without alignments)
699.249 Million cell updates/sec

Title: US-09-991-681-28

Perfect score: 41
Sequence: 1 EDDERSTYSSQCCSEDEDI.....EETAQVSPPRGKERQWRAR 41

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 451899 seqs, 118759770 residues

Word size: 0

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

Published Applications_AA:*
1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubppaa/PTCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubppaa/PTCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubppaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubppaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubppaa/US10C_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	ID	Description
1	7	17.1	898	11 US-09-933-261-5	Sequence 5, Appl1
2	7	17.1	898	15 US-10-256-702-5	Sequence 5, Appl1
3	7	17.1	1404	9 US-09-811-045A-1	Sequence 1, Appl1
4	6	14.6	9	9 US-09-045-020-9	Sequence 9, Appl1
5	6	14.6	16	15 US-10-225-567A-1148	Sequence 1148, Ap
6	6	14.6	22	10 US-09-896-888A-35	Sequence 35, Appl
7	6	14.6	22	10 US-09-896-888A-33	Sequence 33, Appl
8	6	14.6	27	10 US-09-896-888A-29	Sequence 29, Appl
9	6	14.6	28	10 US-09-896-888A-31	Sequence 31, Appl
10	6	14.6	43	10 US-09-896-888A-27	Sequence 27, Appl
11	6	14.6	80	11 US-09-910-082A-310	Sequence 310, App
12	6	14.6	199	14 US-10-027-806-22	Sequence 22, Appl
13	6	14.6	199	14 US-10-034-623-22	Sequence 22, Appl
14	6	14.6	199	15 US-10-027-801-22	Sequence 22, Appl
15	6	14.6	232	10 US-09-738-626-4081	Sequence 4081, Ap

16	6	14.6	252	16 US-10-226-294-32	Sequence 32, Appl
17	6	14.6	282	14 US-10-136-253-7	Sequence 7, Appl
18	6	14.6	339	15 US-10-156-761-9366	Sequence 9366, Ap
19	6	14.6	340	11 US-09-860-836B-7	Sequence 7, Appl1
20	6	14.6	358	14 US-10-001-835-173	Sequence 173, App
21	6	14.6	360	15 US-10-171-311-127	Sequence 127, App
22	6	14.6	374	15 US-10-233-448-8	Sequence 8, Appl1
23	6	14.6	374	15 US-10-106-698-4458	Sequence 4458, Ap
24	6	14.6	406	9 US-09-810-264-12	Sequence 12, Appl
25	6	14.6	413	9 US-09-810-264-14	Sequence 14, Appl
26	6	14.6	454	10 US-09-738-626-6249	Sequence 6249, Ap
27	6	14.6	485	15 US-10-091-752A-14	Sequence 14, Appl
28	6	14.6	512	9 US-09-962-290-2	Sequence 2, Appl1
29	6	14.6	532	15 US-10-225-567A-195	Sequence 195, Appl
30	6	14.6	532	15 US-10-225-567A-2	Sequence 2, Appl1
31	6	14.6	551	14 US-10-005-057A-2	Sequence 2, Appl1
32	6	14.6	567	15 US-10-023-282-573	Sequence 573, App
33	6	14.6	569	14 US-10-014-137-3	Sequence 3, Appl1
34	6	14.6	581	10 US-09-738-626-4914	Sequence 4914, Ap
35	6	14.6	597	15 US-10-128-714-8596	Sequence 8596, Ap
36	6	14.6	604	15 US-10-072-036-41	Sequence 59, Appl
37	6	14.6	605	15 US-10-072-036-41	Sequence 41, Appl
38	6	14.6	649	15 US-10-128-714-3531	Sequence 3531, Ap
39	6	14.6	649	15 US-10-128-714-3531	Sequence 8531, Ap
40	6	14.6	705	15 US-10-154-386-2	Sequence 2, Appl1
41	6	14.6	719	15 US-10-133-642-3	Sequence 3, Appl1
42	6	14.6	891	14 US-10-028-056-4	Sequence 4, Appl1
43	6	14.6	902	14 US-10-043-418-4	Sequence 4, Appl1
44	6	14.6	924	14 US-10-028-056-5	Sequence 5, Appl1
45	6	14.6	1174	15 US-10-205-841-40	Sequence 40, Appl

ALIGNMENTS

RESULT 1
US-09-933-261-5
Sequence 5, Application US/09933261
Publication No. US20030040046A1
GENERAL INFORMATION:
APPLICANT: Tessier-Lavigne, Marc
Leonardo, E. David
Hink, Lindsay
Masu, Masayuki
Kazuko, Keino-Masu
TITLE OF INVENTION: Netrin Receptors
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/933,261
FILING DATE: 20-Aug-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/808,982
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UC96-217
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342

INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 898 amino acids
TYPE: amino acid
STRANDEDNESS: No. US20030040046A1 Relevant
TOPOLOGY: No. US20030040046A1 Relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-933-261-5

Query Match 17.1%; Score 7; DB 11; Length 898;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 ERSTDSS 10
Db 90 ERSTDSS 96

RESULT 2
US-10-256-702-5
Sequence 5, Application US/10256702
Publication No. US20030059859A1
GENERAL INFORMATION:
APPLICANT: Tessier-Lavigne, Marc
Leonardo, E. David
Hink, Lindsay
Masu, Masayuki
Kazuko, Keino-Masu
TITLE OF INVENTION: Nectin Receptors
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSER: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/256,702
FILING DATE: 27-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/933,261
FILING DATE: 20-Aug-2001
APPLICATION NUMBER: 08/808,982
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UC96-217
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 898 amino acids
TYPE: amino acid
STRANDEDNESS: No. US20030059859A1 Relevant
TOPOLOGY: No. US20030059859A1 Relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-256-702-5

Query Match 17.1%; Score 7; DB 15; Length 898;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 ERSTDSS 10
Db 90 ERSTDSS 96

RESULT 3
US-09-811-045A-1
Sequence 1, Application US/09811045A
Patent No. US20020035080A1
GENERAL INFORMATION:
APPLICANT: Scott, Robert E.
TITLE OF INVENTION: cDNA encoding P2P proteins and use of P2P cDNA-
TITLE OF INVENTION: derived antibodies and antisense reagents
TITLE OF INVENTION: in determining the proliferative potential of
TITLE OF INVENTION: normal, abnormal and cancer cells in animals
FILE REFERENCE: D6386D
CURRENT APPLICATION NUMBER: US/09/811,045A
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: US 08/801,308
PRIOR FILING DATE: 1997-02-18
NUMBER OF SEQ ID NOS: 4
SEQ ID NO 1
LENGTH: 1404
TYPE: PRT
ORGANISM: Unknown
FEATURE:
NAME/KEY: PEPTIDE
OTHER INFORMATION: P2P polypeptide
US-09-811-045A-1

Query Match 17.1%; Score 7; DB 9; Length 1404;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 28 SPRGKE 34
Db 1288 SPRGKE 1294

RESULT 4
US-09-045-020-9
Sequence 9, Application US/09045020
Patent No. US20020025942A1
GENERAL INFORMATION:
APPLICANT: Ingram, Vernon M., Roder, Hanno M.
TITLE OF INVENTION: No. US20020025942A1 Tau/Neurofilament Protein
Kinases
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSER: Massachusetts Institute of Technology
STREET: 28 Carleton Street
CITY: Cambridge
STATE: Massachusetts
COUNTRY: United States of America
ZIP: 02142
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch
COMPUTER: IBM-compatible
OPERATING SYSTEM: MS-DOS Version 3.3
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/045,020
FILING DATE: 20-Mar-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/480,793
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: 07/912,293
FILING DATE: July 10, 1992
APPLICATION NUMBER: 07/742,880
FILING DATE: 9-AUG-1991
ATTORNEY/AGENT INFORMATION:

```

; NAME: Gates, Edward R.
; REGISTRATION NUMBER: 31,616
; REFERENCE/DOCKET NUMBER: M0656/7008
; INFORMATION FOR SEQ ID NO: 9:
;     SEQUENCE CHARACTERISTICS:
;         LENGTH: 9 amino acids
;         TYPE: amino acid
;         TOPOLOGY: linear
;     MOLECULE TYPE: peptide
;     SEQUENCE DESCRIPTION: SEQ ID NO: 9
US-09-045-020-9

Query Match          14.6%; Score 6; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      20 IFEETA 25
        |||||
        3 IFEETA 8

RESULT 5
US-10-225-567A-1148
; Sequence 1148, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
;     APPLICANT: Lifespan Biosciences
;     APPLICANT: Brown, Joseph P.
;     APPLICANT: Burnet, Glenn C.
;     APPLICANT: Roush, Christine L.
;     TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
;     FILE REFERENCE: 1920-4-4
;     CURRENT APPLICATION NUMBER: US/10/225,567A
;     CURRENT FILING DATE: 2001-12-19
;     PRIOR APPLICATION NUMBER: 60/257,144
;     PRIOR FILING DATE: 2000-12-19
;     NUMBER OF SEQ ID NOS: 2292
;     SOFTWARE: PatentIn version 3.1
;     SEQ ID NO 1148
;     LENGTH: 16
;     TYPE: PRT
;     ORGANISM: Homo sapiens
US-10-225-567A-1148

Query Match          14.6%; Score 6; DB 15; Length 16;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      14 SSEDED 19
        |||||
        6 SSEDED 11

RESULT 6
US-09-896-888A-35
; Sequence 35, Application US/09896888A
; Patent No. US20020116723A1
; GENERAL INFORMATION:
;     APPLICANT: The University of British Columbia
;     TITLE OF INVENTION: Insect Expression Vectors
;     FILE REFERENCE: 80021-44
;     CURRENT APPLICATION NUMBER: US/09/896,888A
;     CURRENT FILING DATE: 2001-06-29
;     PRIOR APPLICATION NUMBER: US/09/048,911
;     PRIOR FILING DATE: 1998-03-26
;     PRIOR APPLICATION NUMBER: 60/049,946
;     PRIOR FILING DATE: 1997-03-27
;     NUMBER OF SEQ ID NOS: 50
;     SOFTWARE: PatentIn Ver. 2.0
;     SEQ ID NO 35
;     LENGTH: 22
;     TYPE: PRT
;     ORGANISM: Artificial Sequence
```

```

; FEATURE:
;     OTHER INFORMATION:
US-09-896-888A-35

Query Match          14.6%; Score 6; DB 10; Length 22;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      9 SSQOCS 14
        |||||
        10 SSQOCS 15

RESULT 7
US-09-896-888A-33
; Sequence 33, Application US/09896888A
; Patent No. US20020116723A1
; GENERAL INFORMATION:
;     APPLICANT: The University of British Columbia
;     TITLE OF INVENTION: Insect Expression Vectors
;     FILE REFERENCE: 80021-44
;     CURRENT APPLICATION NUMBER: US/09/896,888A
;     CURRENT FILING DATE: 2001-06-29
;     PRIOR APPLICATION NUMBER: US/09/048,911
;     PRIOR FILING DATE: 1998-03-26
;     PRIOR APPLICATION NUMBER: 60/049,946
;     PRIOR FILING DATE: 1997-03-27
;     NUMBER OF SEQ ID NOS: 50
;     SOFTWARE: PatentIn Ver. 2.0
;     SEQ ID NO 33
;     LENGTH: 23
;     TYPE: PRT
;     ORGANISM: Artificial Sequence
;     FEATURE:
;     OTHER INFORMATION:
US-09-896-888A-33

Query Match          14.6%; Score 6; DB 10; Length 23;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      9 SSQOCS 14
        |||||
        10 SSQOCS 15

RESULT 8
US-09-896-888A-29
; Sequence 29, Application US/09896888A
; Patent No. US20020116723A1
; GENERAL INFORMATION:
;     APPLICANT: The University of British Columbia
;     TITLE OF INVENTION: Insect Expression Vectors
;     FILE REFERENCE: 80021-44
;     CURRENT APPLICATION NUMBER: US/09/896,888A
;     CURRENT FILING DATE: 2001-06-29
;     PRIOR APPLICATION NUMBER: US/09/048,911
;     PRIOR FILING DATE: 1998-03-26
;     PRIOR APPLICATION NUMBER: 60/049,946
;     PRIOR FILING DATE: 1997-03-27
;     NUMBER OF SEQ ID NOS: 50
;     SOFTWARE: PatentIn Ver. 2.0
;     SEQ ID NO 29
;     LENGTH: 27
;     TYPE: PRT
;     ORGANISM: Artificial Sequence
;     FEATURE:
;     OTHER INFORMATION:
US-09-896-888A-29

Query Match          14.6%; Score 6; DB 10; Length 27;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 9 SSQOCS 14
 |||||
 Db 10 SSQOCS 15

RESULT 9
 US-09-896-888A-31
 ; Sequence 31, Application US/09896888A
 ; Patent No. US20020116723A1
 ; GENERAL INFORMATION:
 ; APPLICANT: The University of British Columbia
 ; TITLE OF INVENTION: Insect Expression Vectors
 ; FILE REFERENCE: 80021-44
 ; CURRENT APPLICATION NUMBER: US/09/896,888A
 ; CURRENT FILING DATE: 2001-06-29
 ; PRIOR APPLICATION NUMBER: US/09/048,911
 ; PRIOR FILING DATE: 1998-03-26
 ; PRIOR APPLICATION NUMBER: 60/049,946
 ; PRIOR FILING DATE: 1997-03-27
 ; NUMBER OF SEQ ID NOS: 50
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 31
 ; LENGTH: 28
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: :
 US-09-896-888A-31

Query Match 14.6%; Score 6; DB 10; Length 28;
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 SSQOCS 14
 |||||
 Db 10 SSQOCS 15

RESULT 10
 US-09-896-888A-27
 ; Sequence 27, Application US/09896888A
 ; Patent No. US20020116723A1
 ; GENERAL INFORMATION:
 ; APPLICANT: The University of British Columbia
 ; TITLE OF INVENTION: Insect Expression Vectors
 ; FILE REFERENCE: 80021-44
 ; CURRENT APPLICATION NUMBER: US/09/896,888A
 ; CURRENT FILING DATE: 2001-06-29
 ; PRIOR APPLICATION NUMBER: US/09/048,911
 ; PRIOR FILING DATE: 1998-03-26
 ; PRIOR APPLICATION NUMBER: 60/049,946
 ; PRIOR FILING DATE: 1997-03-27
 ; NUMBER OF SEQ ID NOS: 50
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 27
 ; LENGTH: 43
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: :
 US-09-896-888A-27

Query Match 14.6%; Score 6; DB 10; Length 43;
 Best Local Similarity 100.0%; Pred. No. 37;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 SSQOCS 14
 |||||
 Db 10 SSQOCS 15

RESULT 11

US-09-910-082A-310
 ; Sequence 310, Application US/09910082A
 ; Publication No. US20030119731A1
 ; GENERAL INFORMATION:
 ; APPLICANT: University of Utah Research Foundation
 ; APPLICANT: Cognetix, Inc.
 ; APPLICANT: Olivera, Baldomero M.
 ; APPLICANT: McIntosh, J. Michael
 ; APPLICANT: Watkins, Maren
 ; APPLICANT: Garrett, James E.
 ; APPLICANT: Shon, Ki-Joon
 ; APPLICANT: Jacobsen, Richard
 ; APPLICANT: Jones, Robert M.
 ; APPLICANT: Cartier, G. Edward
 ; TITLE OF INVENTION: Omega-Conopeptides
 ; FILE REFERENCE: 2314-241
 ; CURRENT APPLICATION NUMBER: US/09/910,082A
 ; CURRENT FILING DATE: 2001-07-23
 ; PRIOR APPLICATION NUMBER: US 60/219,616
 ; PRIOR FILING DATE: 2000-07-21
 ; PRIOR APPLICATION NUMBER: US 60/265,888
 ; PRIOR FILING DATE: 2001-02-05
 ; NUMBER OF SEQ ID NOS: 413
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 310
 ; LENGTH: 80
 ; TYPE: PRT
 ; ORGANISM: Conus ratus
 US-09-910-082A-310

Query Match 14.6%; Score 6; DB 11; Length 80;
 Best Local Similarity 100.0%; Pred. No. 63;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 RGEKER 36
 |||||
 Db 27 RGEKER 32

RESULT 12
 US-10-027-806-22
 ; Sequence 22, Application US/10027806
 ; Publication No. US20020160476A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Swanson, Ronald V.
 ; APPLICANT: Feldman, Robert A.
 ; APPLICANT: Schleper, Christa
 ; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
 ; FILE REFERENCE: DCOIP.002A
 ; CURRENT APPLICATION NUMBER: US/10/027,806
 ; CURRENT FILING DATE: 2001-12-21
 ; PRIOR APPLICATION NUMBER: 09/408,020
 ; PRIOR FILING DATE: EARLIER APPLICATION NUMBER: 09/408,020
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-29
 ; NUMBER OF SEQ ID NOS: 123
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 22
 ; LENGTH: 199
 ; TYPE: PRT
 ; ORGANISM: Cenarchaeum symbiosum
 US-10-027-806-22

Query Match 14.6%; Score 6; DB 14; Length 199;
 Best Local Similarity 100.0%; Pred. No. 1,4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 DEDIFE 22
 |||||
 Db 87 DEDIFE 92

RESULT 13
 US-10-034-623-22
 ; Sequence 22, Application US/10034623

```
; Publication No. US20020199365A1
; GENERAL INFORMATION:
; APPLICANT: Swanson, Ronald V.
; APPLICANT: Feldman, Robert A.
; APPLICANT: Schleper, Christa
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
; FILE REFERENCE: DORP.002A
; CURRENT APPLICATION NUMBER: US/10/034,623
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/408,020
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: 60/102,294
; PRIOR FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 22
; LENGTH: 199
; TYPE: PRT
; ORGANISM: Cenarchaeum symbiosum
US-10-034-623-22
```

```
Query Match          14.6%; Score 6; DB 14; Length 199;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      17 DEDIFE 22
        |||||
Db      87 DEDIFE 92
```

```
RESULT 14
US-10-027-801-22
; Sequence 22, Application US/10027801
; Publication No. US20030054364A1
; GENERAL INFORMATION:
; APPLICANT: Swanson, Ronald V.
; APPLICANT: Feldman, Robert A.
; APPLICANT: Schleper, Christa
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
; FILE REFERENCE: DORP.002A
; CURRENT APPLICATION NUMBER: US/10/027,801
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-29
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 22
; LENGTH: 199
; TYPE: PRT
; ORGANISM: Cenarchaeum symbiosum
US-10-027-801-22
```

```
Query Match          14.6%; Score 6; DB 15; Length 199;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      17 DEDIFE 22
        |||||
Db      87 DEDIFE 92
```

```
RESULT 15
US-09-738-626-4081
; Sequence 4081, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
```

```
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4081
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4081
```

```
Query Match          14.6%; Score 6; DB 10; Length 232;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      3 DERSTD 8
        |||||
Db      27 DERSTD 32
```

```
Search completed: July 24, 2003, 12:48:27
Job time : 8.9634 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 12:21:29 ; Search time 61.7701 Seconds

(without alignments)
577.701 Million cell updates/sec

Title: US-09-991-681-28

Perfect score: 41
Sequence: 1 EDDERSTDSQOCSEDEDI.....ETAAQVSPRGRKRWRRAR 41

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 5580241 seqs, 870357830 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5580241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep:*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep:*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/1/paa/US097A_COMB.pep:*
22: /cgn2_6/ptodata/1/paa/US097B_COMB.pep:*
23: /cgn2_6/ptodata/1/paa/US098_COMB.pep:*
24: /cgn2_6/ptodata/1/paa/US099A_COMB.pep:*
25: /cgn2_6/ptodata/1/paa/US099B_COMB.pep:*
26: /cgn2_6/ptodata/1/paa/US100_COMB.pep:*
27: /cgn2_6/ptodata/1/paa/US101_COMB.pep:*
28: /cgn2_6/ptodata/1/paa/US102_COMB.pep:*
29: /cgn2_6/ptodata/1/paa/US103_COMB.pep:*
30: /cgn2_6/ptodata/1/paa/US104_COMB.pep:*
31: /cgn2_6/ptodata/1/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	100.0	41	US-08-842-385-7	Sequence 7, Appl1
2	41	100.0	41	US-09-991-681-28	Sequence 28, Appl1

3	41	100.0	467	12	US-08-842-385-6	Sequence 6, Appl1
4	41	100.0	518	25	US-09-991-681-27	Sequence 27, Appl1
5	41	100.0	1770	1	PCT-US03-01943-44	Sequence 44, Appl1
6	41	100.0	1770	27	US-10-144-198-44	Sequence 44, Appl1
7	41	100.0	1807	1	PCT-US01-08631-40087	Sequence 40087, A
8	41	100.0	1839	1	PCT-US01-42950-495	Sequence 495, App
9	41	100.0	1839	30	US-10-416-993-495	Sequence 495, App
10	41	100.0	1872	1	PCT-US03-04508-32	Sequence 32, Appl1
11	41	100.0	1982	1	PCT-US01-08631-40090	Sequence 40090, A
12	41	100.0	2221	1	PCT-US03-01943-30	Sequence 30, Appl1
13	41	100.0	2221	27	US-10-144-198-30	Sequence 30, Appl1
14	13	31.7	192	20	US-09-623-791-87	Sequence 87, Appl1
15	13	31.7	192	20	US-09-623-791A-87	Sequence 87, Appl1
16	13	31.7	192	27	US-10-131-487A-87	Sequence 87, Appl1
17	8	19.5	52	30	US-10-424-599-265455	Sequence 265455, A
18	8	19.5	58	30	US-10-424-599-265456	Sequence 265456, A
19	8	19.5	101	30	US-10-437-963-203739	Sequence 203739, A
20	8	19.5	685	30	US-10-424-599-265457	Sequence 265457, A
21	8	19.5	1023	30	US-10-437-963-136996	Sequence 136996, A
22	7	17.1	54	30	US-10-437-963-156112	Sequence 156112, A
23	7	17.1	60	20	US-09-617-682A-5924	Sequence 5924, Ap
24	7	17.1	80	20	US-09-617-682A-5923	Sequence 5923, Ap
25	7	17.1	87	30	US-10-424-599-279658	Sequence 279658, A
26	7	17.1	101	17	US-09-391-631-812	Sequence 812, App
27	7	17.1	147	22	US-09-758-464-343	Sequence 343, App
28	7	17.1	147	28	US-10-217-550-343	Sequence 343, App
29	7	17.1	202	22	US-09-758-464-350	Sequence 350, App
30	7	17.1	202	28	US-10-217-550-350	Sequence 350, App
31	7	17.1	213	1	PCT-US99-16949-4	Sequence 4, Appl1
32	7	17.1	213	15	US-09-123-020-4	Sequence 4, Appl1
33	7	17.1	213	17	US-09-361-883A-4	Sequence 4, Appl1
34	7	17.1	213	17	US-09-361-883A-4	Sequence 4, Appl1
35	7	17.1	236	16	US-09-248-796-26820	Sequence 26820, A
36	7	17.1	236	27	US-10-179-131-8182	Sequence 8182, Ap
37	7	17.1	236	31	US-60-096-409-26820	Sequence 26820, A
38	7	17.1	309	31	US-60-360-039-22670	Sequence 114, App
39	7	17.1	403	26	US-10-087-192-114	Sequence 3689, Ap
40	7	17.1	403	31	US-60-443-566-3689	Sequence 3689, Ap
41	7	17.1	403	31	US-60-455-444-6967	Sequence 6967, Ap
42	7	17.1	403	31	US-60-465-241-6967	Sequence 6967, Ap
43	7	17.1	431	31	US-60-360-039-20252	Sequence 20252, A
44	7	17.1	592	30	US-10-438-246-10369	Sequence 10369, A
45	7	17.1	610	20	US-09-629-469A-13012	Sequence 13012, A

ALIGNMENTS

RESULT 1
US-08-842-385-7
Sequence 7, Application US/08842385

GENERAL INFORMATION:
APPLICANT: Russell, John
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL FOR DETECTING DISEASE OF THE PROSTATE
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSO for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/842,385
FILING DATE:
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Porombski, Priscilla E
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 6084.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/937-6365
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 41 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: None
US-08-842-385-7

Query Match 100.0%; Score 41; DB 12; Length 41;
Best Local Similarity 100.0%; Pred. No. 4.6e-34;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDERSTDSQCCSSEDEDFEETAOVSPPRGKEKRWRR 41
DB 1 EDDERSTDSQCCSSEDEDFEETAOVSPPRGKEKRWRR 41

RESULT 2
US-09-991-681-28
Sequence 28, Application US/09991681
GENERAL INFORMATION:
APPLICANT: BILLING-MEDEI, PATRICIA
COHEN, MAURICE
COLPITTS, TRACEY L.
FRIEDMAN, PAULA N.
GORDON, JULIAN
GRIMADOS, EDWARD N.
HODGES, STEVEN C.
KLASS, MICHAEL R.
KRATOCHVIL, JON D.
ROBERTS-RAPF, LISA
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/991,681
FILING DATE: 26-Nov-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/065,383
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6084.US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 41 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: None
SEQUENCE DESCRIPTION: SEQ ID NO: 28:
US-09-991-681-28

Query Match 100.0%; Score 41; DB 25; Length 41;
Best Local Similarity 100.0%; Pred. No. 4.6e-34;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDERSTDSQCCSSEDEDFEETAOVSPPRGKEKRWRR 41
DB 1 EDDERSTDSQCCSSEDEDFEETAOVSPPRGKEKRWRR 41

RESULT 3
US-08-842-385-6
Sequence 6, Application US/08842385
GENERAL INFORMATION:
APPLICANT: Russell, John
COLPITTS, TRACEY
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
FOR DETECTING DISEASE OF THE PROSTATE
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/842,385
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Porombski, Priscilla E
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 6084.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/937-6365
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: None
US-08-842-385-6

Query Match 100.0%; Score 41; DB 12; Length 467;
Best Local Similarity 100.0%; Pred. No. 4.8e-34;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDERSTDSQCCSSEDEDFEETAOVSPPRGKEKRWRR 41
DB 133 EDDERSTDSQCCSSEDEDFEETAOVSPPRGKEKRWRR 173

```
RESULT 4
US-09-991-681-27
; Sequence 27, Application US/09991681
; GENERAL INFORMATION:
; APPLICANT: BILLING-MEDEL, PATRICIA
; COHEN, MAURICE
; COLPITTS, TRACEY L.
; FRIEDMAN, PAULA N.
; GORDON, JULIAN
; GRANADOS, EDWARD N.
; HODGES, STEVEN C.
; KLAAS, MICHAEL R.
; KRATOCHVIL, JON D.
; ROBERTS-RAPP, LISA
; TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
; FOR DETECTING DISEASES OF THE PROSTATE
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/991.681
; FILING DATE: 26-Nov-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/065.383
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Becker, Cheryl L.
; REGISTRATION NUMBER: 35,441
; REFERENCE/DOCKET NUMBER: 6084.US.P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847/935-1729
; TELEFAX: 847/938-2623
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 518 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: None
; SEQUENCE DESCRIPTION: SEQ ID NO: 27:
US-09-991-681-27

Query Match 100.0%; Score 41; DB 25; Length 518;
Best Local Similarity 100.0%; Pred. No. 5.3e-33;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EDDERSTDSSQCCSSEDEDIFEEETAQVSPPRGKEKRWRR 41
Db 184 EDDERSTDSSQCCSSEDEDIFEEETAQVSPPRGKEKRWRR 224

RESULT 5
PCT-US03-01943-44
; Sequence 44, Application PC/TUS0301943
; GENERAL INFORMATION:
; APPLICANT: ORIGENE TECHNOLOGIES INC
; TITLE OF INVENTION: CANCER GENES
; FILE REFERENCE: 3U 9U 901 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/01943
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: US 10/054,935
```

```
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: US 60/356,130
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 10/102,946
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: US 10/117,229
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: US 10/144,198
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: US 10/197,824
; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44
; LENGTH: 1770
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-01943-44

Query Match 100.0%; Score 41; DB 1; Length 1770;
Best Local Similarity 100.0%; Pred. No. 1.7e-32;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EDDERSTDSSQCCSSEDEDIFEEETAQVSPPRGKEKRWRR 41
Db 1436 EDDERSTDSSQCCSSEDEDIFEEETAQVSPPRGKEKRWRR 1476

RESULT 6
US-10-144-198-44
; Sequence 44, Application US/10144198
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies Inc
; TITLE OF INVENTION: Regulated Prostate Cance Genes
; FILE REFERENCE: 9U 105 R1
; CURRENT APPLICATION NUMBER: US/10/144,198
; CURRENT FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 44
; LENGTH: 1770
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-198-44

Query Match 100.0%; Score 41; DB 27; Length 1770;
Best Local Similarity 100.0%; Pred. No. 1.7e-32;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EDDERSTDSSQCCSSEDEDIFEEETAQVSPPRGKEKRWRR 41
Db 1436 EDDERSTDSSQCCSSEDEDIFEEETAQVSPPRGKEKRWRR 1476

RESULT 7
PCT-US01-08631-40087
; Sequence 40087, Application PC/TUS0108631
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-049
; CURRENT APPLICATION NUMBER: PCT/US01/08631
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 40087
; LENGTH: 1807
; TYPE: PRT
; ORGANISM: Homo sapiens
```

FEATURE:
NAME/KEY: DOMAIN
LOCATION: (48)..(62)
OTHER INFORMATION: Fusion glycoprotein F0 domain identified by eMATRIX,
OTHER INFORMATION: accession number PF00523D, p-value=7.188e-10, raw score of 11.39
NAME/KEY: DOMAIN
LOCATION: (941)..(950)
OTHER INFORMATION: Helper component proteinase domain identified by Pfam,
OTHER INFORMATION: accession name Peptidase_C6, E-value=0.0056, Pfam score of 7.9
PCT-US01-08631-40087

Query Match 100.0%; Score 41; DB 1; Length 1807;
Best Local Similarity 100.0%; Pred. No. 1.8e-32;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDERSTDSQCCSEDEDFEETAOVSPRGKKEKRWRR 41
DB 1473 EDDERSTDSQCCSEDEDFEETAOVSPRGKKEKRWRR 1513

RESULT 8
PCT-US01-42950-495
Sequence 495, Application PC/TUS0142950
GENERAL INFORMATION:
APPLICANT: Hysseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 21272-096
CURRENT APPLICATION NUMBER: PCT/US01/42950
CURRENT FILING DATE: 2001-11-16
PRIOR APPLICATION NUMBER: 09/714,936
PRIOR FILING DATE: 2000-11-17
NUMBER OF SEQ ID NOS: 682
SOFTWARE: PatentIn version 3.0
SEQ ID NO 495
LENGTH: 1839
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US01-42950-495

Query Match 100.0%; Score 41; DB 1; Length 1839;
Best Local Similarity 100.0%; Pred. No. 1.8e-32;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDERSTDSQCCSEDEDFEETAOVSPRGKKEKRWRR 41
DB 1505 EDDERSTDSQCCSEDEDFEETAOVSPRGKKEKRWRR 1545

RESULT 9
US-10-416-993-495
Sequence 495, Application US/10416993
GENERAL INFORMATION:
APPLICANT: Hysseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 21272-096
CURRENT APPLICATION NUMBER: US/10/416,993
CURRENT FILING DATE: 2003-11-16
PRIOR APPLICATION NUMBER: 09/714,936
PRIOR FILING DATE: 2000-11-17
NUMBER OF SEQ ID NOS: 682
SOFTWARE: PatentIn version 3.0
SEQ ID NO 495
LENGTH: 1839
TYPE: PRT
ORGANISM: Homo sapiens
US-10-416-993-495

Query Match 100.0%; Score 41; DB 30; Length 1839;
Best Local Similarity 100.0%; Pred. No. 1.8e-32;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDERSTDSQCCSEDEDFEETAOVSPRGKKEKRWRR 41

DB 1505 EDDERSTDSQCCSEDEDFEETAOVSPRGKKEKRWRR 1545

RESULT 10
PCT-US03-04508-32
Sequence 32, Application PC/TUS0304508
GENERAL INFORMATION:
APPLICANT: IDEC PHARMACEUTICALS
APPLICANT: GATELY, DENNIS
TITLE OF INVENTION: PROSTATE SPECIFIC GENES AND THE USE THEREOF IN DESIGN
FILE REFERENCE: 037003/0301985
CURRENT APPLICATION NUMBER: PCT/US03/04508
CURRENT FILING DATE: 2003-02-19
PRIOR APPLICATION NUMBER: 60/357,140
PRIOR FILING DATE: 2002-02-19
PRIOR APPLICATION NUMBER: 60/396,082
PRIOR FILING DATE: 2002-07-17
PRIOR APPLICATION NUMBER: 60/386,759
PRIOR FILING DATE: 2002-06-10
NUMBER OF SEQ ID NOS: 93
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 32
LENGTH: 1872
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US03-04508-32

Query Match 100.0%; Score 41; DB 1; Length 1872;
Best Local Similarity 100.0%; Pred. No. 1.8e-32;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDERSTDSQCCSEDEDFEETAOVSPRGKKEKRWRR 41
DB 1538 EDDERSTDSQCCSEDEDFEETAOVSPRGKKEKRWRR 1578

RESULT 11
PCT-US01-08631-40090
Sequence 40090, Application PC/TUS0108631
GENERAL INFORMATION:
APPLICANT: Hysseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 21272-049
CURRENT APPLICATION NUMBER: PCT/US01/08631
CURRENT FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 09/540,217
PRIOR FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: 09/649,167
PRIOR FILING DATE: 2000-08-23
NUMBER OF SEQ ID NOS: 60736
SOFTWARE: Custom
SEQ ID NO 40090
LENGTH: 1982
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (11)..(25)
OTHER INFORMATION: Fusion glycoprotein F0 domain identified by eMATRIX,
OTHER INFORMATION: accession number PF00523D, p-value=7.188e-10, raw score of 11.39
NAME/KEY: DOMAIN
LOCATION: (1065)..(1074)
OTHER INFORMATION: Helper component proteinase domain identified by Pfam,
OTHER INFORMATION: accession name Peptidase_C6, E-value=0.0056, Pfam score of 7.9
PCT-US01-08631-40090

Query Match 100.0%; Score 41; DB 1; Length 1982;
Best Local Similarity 100.0%; Pred. No. 1.9e-32;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDERSTDSQCCSEDEDFEETAOVSPRGKKEKRWRR 41

Db 1587 EDDERSTDSSQCCSEDEDIFEETAAQVSPPRGKEKQRRAR 1627

RESULT 12

PCT-US03-01943-30

Sequence 30, Application PC/TUS0301943

GENERAL INFORMATION:

APPLICANT: ORIGENE TECHNOLOGIES INC

TITLE OF INVENTION: CANCER GENES

FILE REFERENCE: 3U 9U 901 PCT

CURRENT APPLICATION NUMBER: PCT/US03/01943

CURRENT FILING DATE: 2002-01-25

PRIOR APPLICATION NUMBER: US 10/054,935

PRIOR FILING DATE: 2002-01-25

PRIOR APPLICATION NUMBER: US 60/356,130

PRIOR FILING DATE: 2002-02-14

PRIOR APPLICATION NUMBER: US 10/102,946

PRIOR FILING DATE: 2002-03-22

PRIOR APPLICATION NUMBER: US 10/117,229

PRIOR FILING DATE: 2002-04-08

PRIOR APPLICATION NUMBER: US 10/144,198

PRIOR FILING DATE: 2002-05-14

PRIOR APPLICATION NUMBER: US 10/197,824

PRIOR FILING DATE: 2002-07-19

NUMBER OF SEQ ID NOS: 102

SOFTWARE: PatentIn version 3.1

SEQ ID NO 30

LENGTH: 2221

TYPE: PRT

ORGANISM: Homo sapiens

PCT-US03-01943-30

Query Match

Best Local Similarity 100.0%; Score 41; DB 1; Length 2221;

Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EDDERSTDSSQCCSEDEDIFEETAAQVSPPRGKEKQRRAR 41

Db 1587 EDDERSTDSSQCCSEDEDIFEETAAQVSPPRGKEKQRRAR 1927

RESULT 13

US-10-144-198-30

Sequence 30, Application US/10144198

GENERAL INFORMATION:

APPLICANT: Origene Technologies Inc

TITLE OF INVENTION: Regulated Prostate Cance Genes

FILE REFERENCE: 9U 105 R1

CURRENT APPLICATION NUMBER: US/10/144,198

CURRENT FILING DATE: 2002-05-14

NUMBER OF SEQ ID NOS: 44

SOFTWARE: PatentIn version 3.0

SEQ ID NO 30

LENGTH: 2221

TYPE: PRT

ORGANISM: Homo sapiens

US-10-144-198-30

Query Match

Best Local Similarity 100.0%; Score 41; DB 27; Length 2221;

Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EDDERSTDSSQCCSEDEDIFEETAAQVSPPRGKEKQRRAR 41

Db 1587 EDDERSTDSSQCCSEDEDIFEETAAQVSPPRGKEKQRRAR 1927

RESULT 14

US-09-623-791-87

Sequence 87, Application US/09623791

GENERAL INFORMATION:

APPLICANT: SPECHT, THOMAS

APPLICANT: HINZMANN, BERND

APPLICANT: SCHMITT, ARMIN

APPLICANT: SCHMITT, ARMIN

APPLICANT: PILARSKY, CHRISTIAN

APPLICANT: DAHL, EDGAR

APPLICANT: ROSENTHAL, ANDRE

TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM PROSTATE TUMOR TISSUE

FILE REFERENCE: ALBRE 11

CURRENT APPLICATION NUMBER: US/09/623,791

CURRENT FILING DATE: 2001-07-10

PRIOR APPLICATION NUMBER: PCT/DE99/00721

PRIOR FILING DATE: 1999-03-09

NUMBER OF SEQ ID NOS: 201

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 87

LENGTH: 192

TYPE: PRT

ORGANISM: Homo sapiens

US-09-623-791-87

Query Match

Best Local Similarity 100.0%; Score 13; DB 20; Length 192;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EDDERSTDSSQCC 13

Db 180 EDDERSTDSSQCC 192

RESULT 15

US-09-623-791A-87

Sequence 87, Application US/09623791A

GENERAL INFORMATION:

APPLICANT: SPECHT, THOMAS

APPLICANT: HINZMANN, BERND

APPLICANT: SCHMITT, ARMIN

APPLICANT: PILARSKY, CHRISTIAN

APPLICANT: DAHL, EDGAR

APPLICANT: ROSENTHAL, ANDRE

TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM PROSTATE TUMOR TISSUE

FILE REFERENCE: ALBRE 11

CURRENT APPLICATION NUMBER: US/09/623,791A

CURRENT FILING DATE: 2000-09-08

PRIOR APPLICATION NUMBER: PCT/DE99/00721

PRIOR FILING DATE: 1999-03-09

NUMBER OF SEQ ID NOS: 201

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 87

LENGTH: 192

TYPE: PRT

ORGANISM: Homo sapiens

US-09-623-791A-87

Query Match

Best Local Similarity 100.0%; Score 13; DB 20; Length 192;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EDDERSTDSSQCC 13

Db 180 EDDERSTDSSQCC 192

Search completed: July 24, 2003, 12:45:54

Job time : 62.7701 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 12:22:19 ; Search time 1.20059 Seconds
(without alignments)
140.227 Million cell updates/sec

Title: US-09-991-681-28

Perfect score: 41
Sequence: 1 EDERSTDSQOCSSSEDEDI.....ETRAQVSPRGKEXKQWRAR 41

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 41799 seqs, 4106219 residues

Word size: 0

Total number of hits satisfying chosen parameters: 41799

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/2/paa/PCF_NEW_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	100.0	1872	US-10-367-978-32	Sequence 32, Appl
2	14.6	360	1	PCT-US02-18638A-127	Sequence 127, Appl
3	14.6	1129	7	US-60-479-073-17	Sequence 17, Appl
4	14.6	1537	6	US-10-273-573-10709	Sequence 10709, A
5	12.2	16	1	PCT-US03-06553-228	Sequence 228, App
6	12.2	16	6	US-10-378-173-228	Sequence 228, App
7	12.2	19	6	PCT-US03-06553-130	Sequence 130, App
8	12.2	19	6	US-10-378-173-130	Sequence 130, App
9	12.2	21	1	PCT-US03-06553-167	Sequence 167, App
10	12.2	21	6	US-10-378-173-167	Sequence 167, App
11	12.2	45	7	US-60-479-962-461	Sequence 461, App
12	12.2	85	6	US-10-273-573-8009	Sequence 8009, App
13	12.2	90	6	US-10-273-573-8011	Sequence 8011, App
14	12.2	96	6	US-10-308-817-104	Sequence 104, App
15	12.2	107	6	US-10-273-573-5652	Sequence 5652, App
16	12.2	112	5	US-09-784-553C-34	Sequence 34, Appl
17	12.2	112	6	US-10-209-201C-34	Sequence 34, Appl
18	12.2	125	6	US-10-273-573-10391	Sequence 10391, A
19	12.2	126	6	US-10-294-433-333	Sequence 333, App
20	12.2	129	6	US-10-273-573-6201	Sequence 6201, App
21	12.2	140	6	US-10-273-573-10371	Sequence 10371, A
22	12.2	142	6	US-10-273-573-6083	Sequence 6083, App
23	12.2	144	6	US-10-451-554-8	Sequence 8, Appl
24	12.2	151	6	US-10-273-573-7969	Sequence 7969, App
25	12.2	159	6	US-10-273-573-9739	Sequence 9739, App

27	5	12.2	164	6	US-10-294-433-674	Sequence 674, App
28	5	12.2	183	6	US-10-273-573-10372	Sequence 10372, A
29	5	12.2	184	6	US-10-273-573-8825	Sequence 8825, App
30	5	12.2	225	6	US-10-273-573-9769	Sequence 9769, App
31	5	12.2	230	6	US-10-273-573-8649	Sequence 8649, App
32	5	12.2	244	6	US-10-273-573-6235	Sequence 6235, App
33	5	12.2	247	6	US-10-275-595A-18	Sequence 18, Appl
34	5	12.2	260	6	US-10-294-433-420	Sequence 420, App
35	5	12.2	289	6	US-10-273-573-8942	Sequence 8942, App
36	5	12.2	292	5	US-09-903-814A-18	Sequence 18, Appl
37	5	12.2	314	1	PCT-US03-10753-180	Sequence 180, App
38	5	12.2	315	1	PCT-US03-11539-14	Sequence 14, Appl
39	5	12.2	321	6	US-10-273-573-6203	Sequence 6203, App
40	5	12.2	324	1	PCT-US03-03165-2	Sequence 2, Appl
41	5	12.2	329	6	US-10-343-063A-10	Sequence 10, Appl
42	5	12.2	333	5	US-09-290-586A-18	Sequence 18, Appl
43	5	12.2	339	6	US-10-462-022-2	Sequence 2, Appl
44	5	12.2	339	1	PCT-US03-09929-88	Sequence 88, Appl
45	5	12.2	343	1	PCT-US03-09929-82	Sequence 82, Appl

ALIGNMENTS

```
RESULT 1
US-10-367-978-32
; Sequence 32, Application US/10367978
; GENERAL INFORMATION:
; APPLICANT: GATELY, DENNIS
; TITLE OF INVENTION: PROSTATE SPECIFIC GENES AND THE USE THEREOF IN DESIGN
; FILE REFERENCE: 037003-0301988
; CURRENT APPLICATION NUMBER: US/10/367,978
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: 60/357,140
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: 60/396,082
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 60/386,759
; PRIOR FILING DATE: 2002-06-10
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 1872
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-367-978-32

Query Match      100.0%; Score 41; DB 6; Length 1872;
Best Local Similarity 100.0%; Pred. No. 2.9e-35;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1 EDERSTDSQOCSSSEDEDIPEETAQVSPRGKEXKQWRAR 41
Db      1538 EDERSTDSQOCSSSEDEDIPEETAQVSPRGKEXKQWRAR 1578

RESULT 2
PCT-US02-18638A-127
; Sequence 127, Application PC/TUS0218638A
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc. et al.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; FILE REFERENCE: MRI-035PC
; CURRENT APPLICATION NUMBER: PCT/US02/18638A
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
```

PRIOR FILING DATE: 2001-11-14
 ; NUMBER OF SEQ ID NOS: 238
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 127
 ; LENGTH: 360
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 PCT-US02-18638A-127

Query Match 14.6%; Score 6; DB 1; Length 360;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 IFEETA 25
 |||||
 Db 347 IFEETA 352

RESULT 3
 US-60-479-073-17
 ; Sequence 17, Application US/60479073
 ; GENERAL INFORMATION:
 ; APPLICANT: De Wilde, Gert Jules Hector
 ; APPLICANT: Saunders, Michael John Scott
 ; APPLICANT: Logghe, Marc Georges
 ; TITLE OF INVENTION: Amino acid sequences useful for developing compounds for the
 ; TITLE OF INVENTION: Prevention and/or treatment of metabolic diseases and nucleotide
 ; FILE REFERENCE: D00590.70042.US
 ; CURRENT FILING DATE: 2003-06-17
 ; NUMBER OF SEQ ID NOS: 526
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 17
 ; LENGTH: 1129
 ; TYPE: PRT
 ; ORGANISM: Caenorhabditis elegans
 US-60-479-073-17

Query Match 14.6%; Score 6; DB 7; Length 1129;
 Best Local Similarity 100.0%; Pred. No. 37;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 QCSSED 17
 |||||
 Db 7 QCSSED 12

RESULT 4
 US-10-273-573-10709
 ; Sequence 10709, Application US/10273573
 ; GENERAL INFORMATION:
 ; APPLICANT: HySeq, Inc
 ; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
 ; FILE REFERENCE: 21272-066
 ; CURRENT APPLICATION NUMBER: US/10/273,573
 ; CURRENT FILING DATE: 2002-10-18
 ; PRIOR APPLICATION NUMBER: 09/522,929
 ; PRIOR FILING DATE: 2000-04-18
 ; PRIOR APPLICATION NUMBER: 09/770,160
 ; PRIOR FILING DATE: 2001-01-26
 ; NUMBER OF SEQ ID NOS: 10994
 ; SOFTWARE: Custom
 ; SEQ ID NO 10709
 ; LENGTH: 1537
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-273-573-10709

Query Match 14.6%; Score 6; DB 6; Length 1537;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 RGKEXR 36
 |||||
 Db 1062 RGKEXR 1067

RESULT 5
 PCT-US03-06553-228
 ; Sequence 228, Application PC/TUS0306553
 ; GENERAL INFORMATION:
 ; APPLICANT: MDS PROTEOMICS INC.
 ; TITLE OF INVENTION: PHOSPHORYLATED PROTEINS AND USES RELATED THERETO
 ; FILE REFERENCE: MDSP-P01-023
 ; CURRENT APPLICATION NUMBER: PCT/US03/06553
 ; CURRENT FILING DATE: 2003-03-03
 ; PRIOR APPLICATION NUMBER: 60/360787
 ; PRIOR FILING DATE: 2002-03-01
 ; NUMBER OF SEQ ID NOS: 231
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 228
 ; LENGTH: 16
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: phosphorylated peptide
 ; NAME/KEY: MISC_FEATURE
 ; LOCATION: (7)..(7)
 ; OTHER INFORMATION: phosphorylation
 PCT-US03-06553-228

Query Match 12.2%; Score 5; DB 1; Length 16;
 Best Local Similarity 100.0%; Pred. No. 8.8;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 SEDDED 19
 |||||
 Db 7 SEDDED 11

RESULT 6
 US-10-378-173-228
 ; Sequence 228, Application US/10378173
 ; GENERAL INFORMATION:
 ; APPLICANT: Burke et al.
 ; TITLE OF INVENTION: PHOSPHORYLATED PROTEINS AND USES RELATED THERETO
 ; FILE REFERENCE: MDSP-P01-023
 ; CURRENT APPLICATION NUMBER: US/10/378,173
 ; CURRENT FILING DATE: 2003-03-03
 ; PRIOR APPLICATION NUMBER: 60/360787
 ; PRIOR FILING DATE: 2002-03-01
 ; NUMBER OF SEQ ID NOS: 231
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 228
 ; LENGTH: 16
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: phosphorylated peptide
 ; NAME/KEY: MISC_FEATURE
 ; LOCATION: (7)..(7)
 ; OTHER INFORMATION: phosphorylation
 US-10-378-173-228

Query Match 12.2%; Score 5; DB 6; Length 16;
 Best Local Similarity 100.0%; Pred. No. 8.8;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 SEDDED 19
 |||||
 Db 7 SEDDED 11

RESULT 7
PCT-US03-06553-130
; Sequence 130, Application PC/TUS0306553
; GENERAL INFORMATION:
; APPLICANT: MDS PROTEOMICS INC.
; TITLE OF INVENTION: PHOSPHORYLATED PROTEINS AND USES RELATED THERETO
; FILE REFERENCE: MSP-PWO-023
; CURRENT APPLICATION NUMBER: PCT/US03/06553
; CURRENT FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: 60/360787
; PRIOR FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 231
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 130
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: phosphorylated peptide
; NAME/KEY: MISC FEATURE
; LOCATION: (6)..(6)
; FEATURE:
; OTHER INFORMATION: phosphorylation
; NAME/KEY: MISC FEATURE
; LOCATION: (7)..(7)
; OTHER INFORMATION: phosphorylation
PCT-US03-06553-130

Query Match 12.2%; Score 5; DB 1; Length 19;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 SSEDE 18
|||||
Db 6 SSEDE 10

RESULT 8
US-10-378-173-130
; Sequence 130, Application US/10378173
; GENERAL INFORMATION:
; APPLICANT: Burke et al.
; TITLE OF INVENTION: PHOSPHORYLATED PROTEINS AND USES RELATED THERETO
; FILE REFERENCE: MSP-P01-023
; CURRENT APPLICATION NUMBER: US/10/378,173
; CURRENT FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: 60/360787
; PRIOR FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 231
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 130
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: phosphorylated peptide
; NAME/KEY: MISC FEATURE
; LOCATION: (6)..(6)
; OTHER INFORMATION: phosphorylation
; NAME/KEY: MISC FEATURE
; LOCATION: (7)..(7)
; OTHER INFORMATION: phosphorylation
US-10-378-173-130

Query Match 12.2%; Score 5; DB 6; Length 19;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 SSEDE 18
|||||

Db 6 SSEDE 10

RESULT 9
PCT-US03-06553-167
; Sequence 167, Application PC/TUS0306553
; GENERAL INFORMATION:
; APPLICANT: MDS PROTEOMICS INC.
; TITLE OF INVENTION: PHOSPHORYLATED PROTEINS AND USES RELATED THERETO
; FILE REFERENCE: MSP-PWO-023
; CURRENT APPLICATION NUMBER: PCT/US03/06553
; CURRENT FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: 60/360787
; PRIOR FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 231
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 167
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: phosphorylated peptide
; NAME/KEY: MISC FEATURE
; LOCATION: (4)..(4)
; OTHER INFORMATION: phosphorylation
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (6)..(6)
; OTHER INFORMATION: phosphorylation
; NAME/KEY: MISC FEATURE
; LOCATION: (8)..(8)
; OTHER INFORMATION: phosphorylation
PCT-US03-06553-167

Query Match 12.2%; Score 5; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 SSEDE 19
|||||
Db 8 SSEDE 12

RESULT 10
PCT-US03-20887-7
; Sequence 7, Application PC/TUS0320887
; GENERAL INFORMATION:
; APPLICANT: Impact Diagnostics
; TITLE OF INVENTION: Peptides from the E2, E6 and E7 Proteins of Human Papillomavirus
; TITLE OF INVENTION: 18 for Detecting and/or Diagnosing Cervical and Other Human Papi
; FILE REFERENCE: 3352-2-2-1
; CURRENT APPLICATION NUMBER: PCT/US03/20887
; CURRENT FILING DATE: 2003-07-02
; PRIOR APPLICATION NUMBER: US 60/394,172
; PRIOR FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: US (Not yet assigned)
; PRIOR FILING DATE: 2003-07-01
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Derived from the E7 early coding region of HPV 16
PCT-US03-20887-7

Query Match 12.2%; Score 5; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

US-10-308-817-104
 ; Sequence 104, Application US/10308817
 ; GENERAL INFORMATION:
 ; APPLICANT: Rother, Russell
 ; APPLICANT: Wu, Dayang
 ; TITLE OF INVENTION: HYBRID ANTIBODIES
 ; FILE REFERENCE: 1087-37
 ; CURRENT APPLICATION NUMBER: US/10/308,817
 ; CURRENT FILING DATE: 2002-12-03
 ; NUMBER OF SEQ ID NOS: 195
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 104
 ; LENGTH: 96
 ; TYPE: PRT
 ; ORGANISM: human
 US-10-308-817-104

Query Match 12.2%; Score 5; DB 6; Length 96;
 Best Local Similarity 100.0%; Pred. No. 44;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 STDSS 10
 |||||
 Db 89 STDSS 93

Search completed: July 24, 2003, 12:46:19
 Job time : 2.40059 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 12:16:59 ; Search time 4.9224 Seconds

(without alignments)
801.015 Million cell updates/sec

Title: US-09-991-681-28

Perfect score: 41

Sequence: 1 EDDERSTDSSQCCSEDEDI.....ETAAQVSPPRGKRWRRAR 41

Scoring table: OLIGO

Searched: 283308 seqs, 96168682 residues

Word size: 0

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

PIR_76: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7	17.1	221	2 E86425	hypothetical prote
2	7	17.1	266	2 S44954	lmbg protein - Str
3	7	17.1	318	2 T41105	60S ribosomal prot
4	7	17.1	402	2 T12178	hypothetical prote
5	7	17.1	837	2 S64785	hypothetical prote
6	7	17.1	1560	2 T42727	proliferation pote
7	6	14.6	51	2 D83889	hypothetical prote
8	6	14.6	105	2 B44890	E7 protein - human
9	6	14.6	105	2 S36580	E7 protein - human
10	6	14.6	105	2 S36528	E7 protein - human
11	6	14.6	123	1 F60471	E7 protein - hesu
12	6	14.6	123	1 F60471	hypothetical prote
13	6	14.6	123	2 B91201	hypothetical prote
14	6	14.6	144	2 PC7027	aggreitin alpha cha
15	6	14.6	148	2 S48980	hypothetical prote
16	6	14.6	155	2 B70030	conserved hypotnet
17	6	14.6	158	2 D90959	unknown protein en
18	6	14.6	174	2 E75552	hypothetical prote
19	6	14.6	179	2 G95104	ribosomal protein
20	6	14.6	199	2 T13137	hypothetical prote
21	6	14.6	207	2 S21124	hypothetical prote
22	6	14.6	209	1 S48459	B2 protein - carro
23	6	14.6	220	2 T46055	probable dual spec
24	6	14.6	234	2 E70982	probable magnesium
25	6	14.6	236	2 E72404	flagellar export/a
26	6	14.6	264	2 T38795	conserved uncharac
27	6	14.6	265	2 S10316	conserved uncharac
28	6	14.6	265	2 S10316	conserved uncharac
29	6	14.6	266	2 D69776	hypothetical prote

30	6	14.6	271	2 AB0754	conserved hypotnet
31	6	14.6	277	1 Q0BEC2	HKRF protein - hu
32	6	14.6	279	2 T25579	hypothetical prote
33	6	14.6	288	2 T15735	hypothetical prote
34	6	14.6	297	2 H69708	DNA processing Smf
35	6	14.6	300	2 T02452	probable Skp1-like
36	6	14.6	301	2 T17321	hypothetical prote
37	6	14.6	302	2 F85068	N7 like-protein [i
38	6	14.6	315	2 H96598	protein F20N2.8 [i
39	6	14.6	327	2 T36087	probable binding p
40	6	14.6	327	2 A56545	F1P1 protein - yea
41	6	14.6	328	2 A83519	hypothetical prote
42	6	14.6	339	2 T34925	ABC transporter in
43	6	14.6	349	2 T38904	TATA-binding prote
44	6	14.6	353	2 T34059	hypothetical prote
45	6	14.6	358	2 S16444	mitogen-activated

ALIGNMENTS

RESULT 1
E86425
Hypothetical protein T2H7.6 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
C:Accession: E86425
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Hutzar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Malt, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I. ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: E86425
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-221 <STO>
A:Cross-references: GB:AB005172; NID:g11094784; PIDN:AAG29716.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1
Query Match 17.1%; Score 7; DB 2; Length 221;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cy 29 PPRGKER 35
Db 4 PPRGKER 10
RESULT 2
S44954
lmbg protein - Streptomyces lincolnensis
C:Species: Streptomyces lincolnensis
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 22-Oct-1999
C:Accession: S69816; S44954
R:Peckhe, U.; Schmidt, H.; Zhang, H.Z.; Pieperberg, W.
Mol. Microbiol. 16, 1137-1156, 1995
A:Title: Molecular characterization of the lincomycin-production gene cluster of Streptom
A:Reference number: S69805; MUID:96020646; PMID:8577249
A:Accession: S69816
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-266 <PE2>
A:Cross-references: EMBL:X79146; NID:g499194; PIDN:CA55753.1; PID:g487697
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1994
C:Genetics:
A:Gene: lmbg
hypothetical prote

C:Superfamily: Streptomyces lincolnensis lmbg protein

Query Match 17.1%; Score 7; DB 2; Length 266;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 TAQVSP 30
DB 11 TAQVSP 17

RESULT 3

T41105
60s ribosomal protein - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 07-Mar-2003

C:Accession: T41105

R:Putrille, B.; Goffeau, A.; Wood, V.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, September 1998

A:Reference number: Z21964

A:Accession: T41105

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-318 <PUR>

A:Cross-references: EMBL:AL031535; PIDN:CA38631.1; GSPDB:GN00068; SPDB:SPCC16C4.15

A:Experimental source: strain 97zh-; cosmid c16C4

A:Gene: SPDB:SPCC16C4.15

A:Map position: 3

A:Introns: 10/3

C:Superfamily: ribosomal protein l2

Query Match 17.1%; Score 7; DB 2; Length 318;
Best Local Similarity 100.0%; Pred. No. 9.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 PRGKEK 36
DB 312 PRGKEK 318

RESULT 4
T21278
hypothetical protein F22G12.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T21278

R:Lennard, N.

submitted to the EMBL Data Library, March 1997

A:Reference number: Z19400

A:Accession: T21278

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-402 <WIL>

A:Cross-references: EMBL:Z92831; PIDN:CAM07366.1; GSPDB:GN00019; CESP:F22G12.2

A:Experimental source: clone F22G12

C:Genetics:

A:Gene: CESP:F22G12.2

A:Map position: 1

Query Match 17.1%; Score 7; DB 2; Length 402;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 EDIFEET 24
DB 188 EDIFEET 194

C:Species: Saccharomyces cerevisiae

C>Date: 01-Aug-1995 #sequence_revision 24-May-1996 #text_change 19-Apr-2002

C:Accession: S64785

R:Duesterhoeft, A.; Floeth, M.; Heuss-Neitzel, D.; Hilbert, H.; Moestl, D.

submitted to the Protein Sequence Database, May 1996

A:Reference number: S64775

A:Accession: S64785

A:Molecule type: DNA

A:Residues: 1-837 <DUE>

A:Cross-references: EMBL:Z71139; NID:G1360221; PIDN:CAA97483.1; PID:G1360222; MIPS:YLL03

A:Experimental source: strain S288C

C:Genetics:

A:Cross-references: SGD:S0003957

A:Map position: 12L

C:Superfamily: transitional endoplasmic reticulum ATPase; FtsH/SEC18/CDC48-type ATP-bind

C:Keywords: ATP; nucleotide binding; P-loop

F:219-431/Domain: FtsH/SEC18/CDC48-type ATP-binding domain homology <VAT1>

F:246-253/Region: nucleotide-binding motif A (P-loop)

F:547-758/Domain: FtsH/SEC18/CDC48-type ATP-binding domain homology <VAT2>

F:574-581/Region: nucleotide-binding motif A (P-loop)

Query Match 17.1%; Score 7; DB 2; Length 837;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 EDEDFE 22
DB 54 EDEDFE 60

RESULT 6
T42727
proliferation potential-related protein - mouse

C:Species: Mus musculus (house mouse)

C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 02-Sep-2000

C:Accession: T42727

R:Witte, M.M.; Scott, R.F.

submitted to the EMBL Data Library, November 1998

A:Reference number: Z22246

A:Accession: T42727

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1560 <WIT>

A:Cross-references: EMBL:U03913; NID:G3858884; PID:G3858885; PIDN:MAC72432.1

A:Experimental source: strain Balb/C

C:Genetics:

A:Gene: P2P-R

C:Function:

A:Description: involved in hnRNP association and Rb1 binding

C:Superfamily: RING finger homology

F:57-107/Domain: RING finger homology <RRN>

Query Match 17.1%; Score 7; DB 2; Length 1560;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 SPPRKE 34
DB 1271 SPPRKE 1277

RESULT 7
D83889
hypothetical protein BH1916 [imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans

C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001

C:Accession: D83889

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and c

A:Reference number: A83650; MUID:2051582; PMID:11058132

A:Accession: D83889

A:Status: preliminary

A:Molecule type: DNA
A:Residues: 1-51 <STO>
A:Cross-references: GB:AP001513; GB:BA000004; NID:g10174345; PIDN:BAB05635.1; GSPDB:GN00
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH1916

Query Match 14.6%; Score 6; DB 2; Length 51;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 29 PPRGKE 34
DB 41 PPRGKE 46

RESULT 8
B44890
E7 protein - human papillomavirus type 66
C:Species: human papillomavirus type 66
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 22-May-1998
C:Accession: B44890
R:Tawheed, A.R.; Beaudenon, S.; Favre, M.; Orth, G.
J. Clin. Microbiol. 29, 2656-2660, 1991
A:Title: Characterization of human papillomavirus type 66 from an invasive carcinoma of
A:Reference number: A44890; MUID:92129556; PMID:1663515
A:Accession: B44890
A:Molecule type: DNA
A:Residues: 1-105 <TAM>
A>Note: sequence extracted from NCBI backbone (NCBI:78637, NCBI:78639)
C:Superfamily: papillomavirus E7 protein
C:Keywords: DNA binding; early protein; transcription regulation

Query Match 14.6%; Score 6; DB 2; Length 105;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 14 SSEDED 19
DB 31 SSEDED 36

RESULT 9
S36580
E7 protein - human papillomavirus type 56
C:Species: human papillomavirus type 56
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
C:Accession: S36580
R:DeJure, H.; Hofmann, B.
submitted to the EMBL Data Library, August 1993
A:Description: Primer-directed sequencing of human papillomavirus types.
A:Reference number: S36469
A:Accession: S36580
A:Molecule type: DNA
A:Residues: 1-105 <DELU>
A:Cross-references: EMBL:X74483; NID:g397053; PIDN:CAA52597.1; PID:g397055
C:Superfamily: papillomavirus E7 protein
C:Keywords: DNA binding; early protein; transcription regulation

Query Match 14.6%; Score 6; DB 2; Length 105;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 14 SSEDED 19
DB 31 SSEDED 36

RESULT 10
S36528
E7 protein - human papillomavirus type 53
C:Species: human papillomavirus type 53
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999

C:Accession: S36528
R:DeJure, H.; Hofmann, B.
submitted to the EMBL Data Library, August 1993
A:Description: Primer-directed sequencing of human papillomavirus types.
A:Reference number: S36469
A:Accession: S36528
A:Molecule type: DNA
A:Residues: 1-105 <DELU>
A:Cross-references: EMBL:X74483; NID:g397046; PIDN:CAA52592.1; PID:g397048
C:Superfamily: papillomavirus E7 protein
C:Keywords: DNA binding; early protein; transcription regulation

Query Match 14.6%; Score 6; DB 2; Length 105;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 14 SSEDED 19
DB 31 SSEDED 36

RESULT 11
W7MER1
E7 protein - rhesus papillomavirus (type 1)
C:Species: rhesus papillomavirus
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 16-Feb-1997
C:Accession: B38503
R:Ostrow, R.S.; Labresh, K.V.; Faras, A.J.
Virology 181, 424-429, 1991
A:Title: Characterization of the complete RHPV 1 genomic sequence and an integration loc
A:Reference number: A38503; MUID:9135018; PMID:1847267
A:Accession: B38503
A>Status: translation not shown

A:Molecule type: DNA
A:Residues: 1-113 <OST>
A:Cross-references: EMBL:M37717
C:Superfamily: papillomavirus E7 protein
C:Keywords: DNA binding; early protein; transcription regulation

Query Match 14.6%; Score 6; DB 1; Length 113;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 14 SSEDED 19
DB 35 SSEDED 40

RESULT 12
F86047
hypothetical protein Z5129 [imported] - Escherichia coli (strain O157:H7, substrain EDL9
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: F86047
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; DiMantano, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: F86047
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-123 <STO>
A:Cross-references: GB:AE005174; NID:g12518469; PIDN:AA65842.1; GSPDB:GN00145; UMGF:Z51:
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: Z5129

Query Match 14.6%; Score 6; DB 2; Length 123;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 15 SEDEDI 20

Db 45 SEDEDI 50

RESULT 13

B91201

hypothetical protein ECs4578 (imported) - Escherichia coli (strain O157:H7, substrain R1
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C:Accession: B91201
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
A:Reference number: A99629; MUID:2115631; PMID:11258796
A:Accession: B91201
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-123 <HAV>
A:Cross-references: GB:BA000007; PDB:BA038001.1; PID:g13364053; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain R1MD 0509952
C:Genetics:
A:Gene: ECs4578

Query Match 14.6%; Score 6; DB 2; Length 123;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 SEDEDI 20
Db 45 SEDEDI 50

RESULT 14

PC7027

aggrexin alpha chain - Malaysian pit viper (fragment)
C:Species: Calloselasma rhodostoma (Malayan pit viper)
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 02-Jun-2000
C:Accession: PC7027
R:Chung, C.H.; Au, L.C.; Huang, T.F.
Biochem. Biophys. Res. Commun. 263, 723-727, 1999
A:Title: Molecular cloning and sequence analysis of aggrexin, a collagen-like platelet a
A:Reference number: PC7027; MUID:99443731; PMID:10512747
A:Accession: PC7027
A:Molecule type: mRNA
A:Residues: 1-144 <CHU>
A:Experimental source: venom gland
C:Superfamily: tetranectin; C-type lectin homology
C:Keywords: disulfide bond; platelet aggregation; venom

Query Match 14.6%; Score 6; DB 2; Length 144;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 OQCSE 16
Db 89 OQCSE 94

RESULT 15

S48980

hypothetical protein YHR136c - yeast (Saccharomyces cerevisiae)
C:Species: Saccharomyces cerevisiae
C:Date: 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 23-Mar-2001
C:Accession: S48980
R:Fulton, L.
submitted to the EMBL Data Library, June 1994
A:Description: The sequence of S. cerevisiae cosmid 9315.
A:Reference number: S48967
A:Accession: S48980
A:Molecule type: DNA
A:Residues: 1-148 <FUL>
A:Cross-references: EMBL:U10398; NID:g551328; PDB:AAB68421.1; PID:g551329; MIPS:YHR136c

C:Genetics:
A:Gene: SGD:SP12
A:Cross-references: SGD:S0001178; MIPS:YHR136c
A:Map position: 8R
C:Superfamily: Saccharomyces cerevisiae hypothetical protein YHR136c

Query Match 14.6%; Score 6; DB 2; Length 148;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 DEDIFE 22
Db 135 DEDIFE 140

Search completed: July 24, 2003, 12:27:17
Job time : 10.9224 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 12:15:23 ; Search time 2.70132 Seconds
(without alignments)
713.761 Million cell updates/sec

Title: US-09-991-681-28

Perfect score: 41
Sequence: 1 EDDERSTYDSQQCSSEDEDI.....ETTAQVSPRKEKRWRRAR 41

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size: 0

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7	17.1	200	1 CANK_COTJA	Q08855 coturnix co
2	7	17.1	837	1 YL34_YEAST	Q07844 saccharomyc
3	6	14.6	105	1 VET7_HPV53	P36832 human papil
4	6	14.6	105	1 VET7_HPV56	P36833 human papil
5	6	14.6	105	1 VET7_HPV66	Q08956 human papil
6	6	14.6	113	1 VET7_HPV1	P22161 rhesus papi
7	6	14.6	131	1 ABA1_TRIAB	P81111 trimeresuru
8	6	14.6	132	1 ABA2_TRIAB	P81112 trimeresuru
9	6	14.6	134	1 ABA2_TRIAB	P81112 trimeresuru
10	6	14.6	148	1 YHT6_YEAST	P37839 saccharomyc
11	6	14.6	207	1 B2_DAUCA	P37707 dancus caro
12	6	14.6	209	1 YL13_YEAST	P40479 saccharomyc
13	6	14.6	264	1 Y554_SCHPO	O14171 schizosacch
14	6	14.6	265	1 Y554_SCHPO	P18051 bacillus am
15	6	14.6	266	1 YDDK_BACSU	P96648 bacillus su
16	6	14.6	271	1 MGPB_SALTY	O89681 salmonelel
17	6	14.6	277	1 J11_HCMVA	P09711 human cytom
18	6	14.6	282	1 DAA4_STANA	P56694 staphylococ
19	6	14.6	288	1 Y005_CAEEL	Q09265 caenorhabdi
20	6	14.6	297	1 SMF_BACSU	P39813 bacillus su
21	6	14.6	301	1 PHLF_HUMAN	Q13371 homo sapien
22	6	14.6	313	1 IRE1_CHICK	Q09876 gallus gall
23	6	14.6	327	1 F1P1_YEAST	P45976 saccharomyc
24	6	14.6	349	1 NTRK_RAT	Q94766 ratiscus norv
25	6	14.6	349	1 Y12D6_HUMAN	Q15545 homo sapien
26	6	14.6	358	1 MKO1_MOUSE	P27703 mus musculu
27	6	14.6	360	1 MKO1_BOVIN	P46196 bos taurus
28	6	14.6	360	1 MKO1_HUMAN	P28482 homo sapien
29	6	14.6	361	1 MKO1_XENLA	P26696 xenopus lae
30	6	14.6	365	1 DCN3_ARATH	Q92886 arabidopsis
31	6	14.6	367	1 DCN3_BRAUJ	Q96dm8 brassica ju
32	6	14.6	395	1 ZAAA_PEA	P36875 pisum sativ
33	6	14.6	426	1 CG2A_PATVU	P24861 patella vul

34	6	14.6	447	1 GNT1_MOUSE	P27808 mus musculu
35	6	14.6	447	1 GNT1_RABIT	P27115 oryctolagus
36	6	14.6	447	1 GNT1_RAT	Q09325 ratiscus norv
37	6	14.6	458	1 YNE1_CAEEL	P30640 caenorhabdi
38	6	14.6	489	1 YNE1_SCHPO	Q09818 schizosacch
39	6	14.6	490	1 CYKL_DROME	P17971 drosophila
40	6	14.6	492	1 TOM1_HUMAN	O60784 homo sapien
41	6	14.6	507	1 YIK4_YEAST	P40486 saccharomyc
42	6	14.6	512	1 FLO1_MOUSE	P41438 mus musculu
43	6	14.6	512	1 ACMS_HUMAN	P08912 homo sapien
44	6	14.6	532	1 ACMS_MACMU	P56480 macaca mula
45	6	14.6	569	1 GCL1_DROME	Q01820 drosophila

ALIGNMENTS

RESULT 1	CANK_COTJA	STANDARD;	PRT;	200 AA.
ID	CANK_COTJA			
AC	Q08855;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	01-NOV-1997 (Rel. 35, Last annotation update)			
DE	Putative calcium-activated potassium channel beta subunit.			
OS	Coturnix coturnix japonica (Japanese quail).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;			
OC	Coturnix.			
OX	NCBI_TaxID=93934;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=97224079; PubMed=9070660;			
RA	Oberst C., Weiskirchen R., Hartl M., Bister K.;			
RT	"Suppression in transformed avian fibroblasts of a gene (CO6)			
RT	encoding a membrane protein related to mammalian potassium channel			
RT	regulatory subunits."			
RL	Oncogene 14:1109-1116(1997).			
CC	-1- FUNCTION: POTASSIUM CHANNEL PROTEIN WHICH MAY MODULATE THE			
CC	PROPERTIES OF THE PORE-FORMING ALPHA SUBUNIT (BY SIMILARITY).			
CC	-1- SUBUNIT: THE CALCIUM-ACTIVATED POTASSIUM CHANNEL IS COMPOSED OF			
CC	AT LEAST TWO SUBUNITS: A PORE-FORMING ALPHA SUBUNIT AND A			
CC	REGULATORY BETA SUBUNIT (BY SIMILARITY).			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; U67865; AAC26967.1; -			
DR	InterPro; IPR003930; BK_channel_beta.			
DR	Pfam; PF03185; CankB_1			
KH	Ionic channel; Transmembrane; Glycoprotein.			
FT	TRANSMEM 20 40			
FT	TRANSMEM 167 187			
FT	CARBOHYD 89 89			
FT	SEQUENCE 200 AA; 22677 MW; B7A612D89737711F CRC64;			
SQ				
Query Match	17.1%; Score 7; DB 1; Length 200;			
Best Local Similarity	100.0%; Pred. No. 3;			
Matches	7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Oy	15 SEDSDIF 21			
Db	73 SEDSDIF 79			
RESULT 2				

```

YL34_YEAST
ID YL34_YEAST STANDARD; PRT; 837 AA.
AC 007844;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein YL034C.
GN YL034C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=5288C / AB972;
RX MEDLINE=9713267; PubMed=9169871;
RA Johnston M., Hillier L., Riles L., Dubois E., Duesterhoeft A.,
RA Benes V., Brueckner M., Delius H., Dubois E., Duesterhoeft A.,
RA Entian K.-D., Floeth M., Goffeau A., Hebling U., Heumann K.,
RA Heuss-Nietzel D., Hilbert H., Hilger F., Kleine K., Koetter P.,
RA Louis E.J., Messenguy F., Mewes H.-W., Miosga T., Moesli D.,
RA Mueller-Auer S., Neutwich U., Obermayer B., Piravandi E., Pohl T.M.,
RA Portetelle D., Purnelle B., Reumann S., Rieger M., Rinke M., Rose M.,
RA Scharte M., Scherrens B., Scholler P., Schwager C., Schwarz S.,
RA Underwood A.P., Urrestarazu L.A., Vandenbol M., Verhasselt P.,
RA Vierderdele F., Voet M., Volckaert G., Voss H., Wambutt R., Medler E.,
RA Wedler H., Zimmermann F.K., Zollner A., Hanl U., Hohnsbeil U.D.,
RA "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";
RT Nature 387:87-90(1997).
RL Nature 387:87-90(1997).
CC -1 SIMILARITY: BELONGS TO THE AAA FAMILY OF ATPASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, Z73139; CAA97483.1; -.
CC DR PIR, S64785; S64785.
CC DR SGD, S0003957; YL034C.
CC DR InterPro: IPR003593; AAA_ATPase.
CC DR InterPro: IPR003959; AAA_ATPase_cent.
CC DR InterPro: IPR003960; AAA_sub.
CC DR Pfam, PF00004; AAA; 2.
CC DR SMART, SM00382; AAA; 1.
CC DR PROSITE, PS00674; AAA; 1.
CC DR Hypothetical protein; ATP-binding; Repeat.
CC NP_BIND 246 253 ATP (POTENTIAL).
CC FT NP_BIND 574 581 ATP (POTENTIAL).
CC FT SEQUENCE 837 AA; 93069 MW; 133528B9D0987103 CRC64;
SQ

```

```

OC Papillomavirus.
OX NCBI_TaxID=10619;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94265501; PubMed=8205838;
RA Delius H., Hofmann B.;
RT "Primer-directed sequencing of human papillomavirus types.";
RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).
CC -1 FUNCTION: E7 PROTEIN HAS BOTH TRANSFORMING AND TRANS-ACTIVATING
CC ACTIVITIES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, X74482; CAA52592.1; -.
CC DR PIR, S36528; S36528.
CC DR InterPro: IPR000148; Papv1_E7.
CC DR Pfam, PF00527; E7; 1.
CC DR Early protein; Transcription regulation; Oncogene;
CC DNA-binding; Trans-acting factor.
CC FT SITE 65 68 C-XX-C MOTIF-1.
CC FT SITE 98 101 C-XX-C MOTIF-2.
CC FT SEQUENCE 105 AA; 12162 MW; DE47E2B8D0C1F7EC CRC64;
SQ

```

Query Match 14.6%; Score 6; DB 1; Length 105;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 14 SEDED 19
DB 31 SEDED 36

```

```

RESULT 4
VE7_HPV56 STANDARD; PRT; 105 AA.
ID VE7_HPV56
AC P36833;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE E7 protein.
GN E7.
OS Human papillomavirus type 56.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10596;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94265501; PubMed=8205838;
RA Delius H., Hofmann B.;
RT "Primer-directed sequencing of human papillomavirus types.";
RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).
CC -1 FUNCTION: E7 PROTEIN HAS BOTH TRANSFORMING AND TRANS-ACTIVATING
CC ACTIVITIES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, X74483; CAA52597.1; -.
CC DR PIR, S36580; S36580.
CC DR InterPro: IPR000148; Papv1_E7.
CC DR Pfam, PF00527; E7; 1.
CC Early protein; Transcription regulation; Oncogene;

```

```

KW DNA-binding; Trans-acting factor.
FT SITE 65 68 C-XX-C MOTIF-1.
FT SITE 98 101 C-XX-C MOTIF-2.
SQ SEQUENCE 105 AA; 11892 MW; 35424B643B4E0183 CRC64;

Query Match
Best Local Similarity 14.6%; Score 6; DB 1; Length 105;
Matches 6; Conservative 0; Pred. No. 18;
Indels 0; Gaps 0;

Oy 14 SSEDDED 19
Db 31 SSEDDED 36

RESULT 5
VE7 HPV66 STANDARD; PRT; 105 AA.
AC Q80356;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE E7 protein.
GN E7.
OS Human papillomavirus type 66.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=37119;
RN [1]
RP SEQUENCE FROM N.A.
RA Delius H.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: E7 PROTEIN HAS BOTH TRANSFORMING AND TRANS-ACTIVATING
CC ACTIVITIES.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; U31794; AAA79500.1; -.
CC PIR; B44890; B44890.
CC InterPro; IPR000148; Papv1_E7.
DR Pfam; PF00527; E7; 1.
KW Early protein; Transcription regulation; Oncogene;
KW DNA-binding; Trans-acting factor.
FT SITE 66 69 C-XX-C MOTIF-1.
FT SITE 98 101 C-XX-C MOTIF-2.
SQ SEQUENCE 105 AA; 11911 MW; C09339F42F62AF6A CRC64;

Query Match
Best Local Similarity 14.6%; Score 6; DB 1; Length 105;
Matches 6; Conservative 0; Pred. No. 18;
Indels 0; Gaps 0;

Oy 14 SSEDDED 19
Db 31 SSEDDED 36

RESULT 6
VE7 RHPV1 STANDARD; PRT; 113 AA.
AC P22161;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE E7 protein.
GN E7.
OS Rheus papillomavirus type 1 (Rhpv 1).
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.

```

```

OX NCBI_TaxID=10570;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91135018; PubMed=1847267;
RA Ostrow R.S., Labresh K.V., Faras A.J.;
RT "Characterization of the complete KnpV 1 genomic sequence and an
RT integration locus from a metastatic tumor.";
RL Virology 181:424-429(1991).
CC -1- FUNCTION: E7 PROTEIN HAS BOTH TRANSFORMING AND TRANS-ACTIVATING
CC ACTIVITIES.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; M60184; AAA79312.1; -.
CC PIR; B38503; W7WLR1.
DR InterPro; IPR000148; Papv1_E7.
DR Pfam; PF00527; E7; 1.
KW Early protein; Transcription regulation; Oncogene;
KW DNA-binding; Trans-acting factor.
FT SITE 73 76 C-XX-C MOTIF-1.
FT SITE 106 109 C-XX-C MOTIF-2.
SQ SEQUENCE 113 AA; 12818 MW; 9C404BDAC128306 CRC64;

Query Match
Best Local Similarity 14.6%; Score 6; DB 1; Length 113;
Matches 6; Conservative 0; Pred. No. 19;
Indels 0; Gaps 0;

Oy 14 SSEDDED 19
Db 35 SSEDDED 40

RESULT 7
AB1 TRIAB STANDARD; PRT; 131 AA.
AC P81111;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Alboagreggin A subunit 1.
OS Trimeresurus albolabris (White-lipped pit viper).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidodonta; Squamata; Scleroglossa; Serpentes; Colubroides;
OC Viperidae; Crotalinae; Trimeresurus.
OX NCBI_TaxID=8765;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=98189535; PubMed=9531050;
RA Kowaleka M.A., Tan L., Holt J.C., Peng M., Karczewski J.,
RA Calvete J.J., Niewiarowski S.;
RT "Alboagreggin A and B. Structure and interaction with human
RT platelets.";
RL Thromb. Haemost. 79:609-613(1998).
CC -1- FUNCTION: Binds to platelet GPIb/IX receptor system and stimulates
CC agglutination.
CC -1- SUBUNIT: Heterotetramer of the subunits 1, 2, 3 and 4,
CC disulfide-linked.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Contains 1 C-type lectin family domain.
DR HSSP; P23806; 11XX.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; Lectin_C; 1.
DR SMART; SM00034; LECTC1.1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW Lectin.

```

```

FT DOMAIN 1 129 C-TYPE LECTIN (LONG FORM).
FT DISULFID 2 13 BY SIMILARITY.
FT DISULFID 30 127 BY SIMILARITY.
FT DISULFID 102 119 BY SIMILARITY.
SQ SEQUENCE 131 AA; 15427 MW; B35695FBF91F6624 CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 1; Length 131;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 QOCSSSE 16
Db 77 QOCSSSE 82

RESULT 8
ABBA TRIAB STANDARD; PRT; 132 AA.
AC P8115;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Albogagregin B alpha subunit.
OS Trimeresurus albolabris (White-lipped pit viper).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidodactylus; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperae; Crotalinae; Trimeresurus.
OC NCB1_TaxID=8765;
OX TISSUE=Venom;
RP SEQUENCE.
RX MEDLINE=98189535; PubMed=9531050;
RA Kowalska M.A., Tan L., Holt J.C., Peng M., Karczewski J.,
RA Calvele J.J., Niewiarowski S.;
RT "Albogagregin A and B. Structure and interaction with human
RT platelets.";
RL Thromb. Haemost. 79:609-613(1998).
CC -1- FUNCTION: Binds to platelet GPIb/IX receptor system and stimulates
CC agglutination.
CC -1- SUBUNIT: Heterodimer of alpha and beta subunits, disulfide-linked.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Contains 1 C-type lectin family domain.
DR HSSP; P23806; 11XX.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C-TYPE LECTIN_1; 1.
DR PROSITE; PS50041; C-TYPE LECTIN_2; 1.
KW Lectin.
FT DOMAIN 1 129 C-TYPE LECTIN (LONG FORM).
FT DISULFID 2 13 BY SIMILARITY.
FT DISULFID 30 127 BY SIMILARITY.
FT DISULFID 79 79 INTERCHAIN (WITH C-75 IN BETA CHAIN)
FT DISULFID (BY SIMILARITY).
FT DISULFID 102 119 BY SIMILARITY.
SQ SEQUENCE 132 AA; 15419 MW; 6FAE64820383F16F CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 1; Length 132;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 QOCSSSE 16
Db 77 QOCSSSE 82

RESULT 9
ABBA TRIAB STANDARD; PRT; 134 AA.
AC P8112;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)

```

```

DE Albogagregin A subunit 2.
OS Trimeresurus albolabris (White-lipped pit viper).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidodactylus; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperae; Crotalinae; Trimeresurus.
OX NCB1_TaxID=8765;
RP SEQUENCE.
RX MEDLINE=98189535; PubMed=9531050;
RA Kowalska M.A., Tan L., Holt J.C., Peng M., Karczewski J.,
RA Calvele J.J., Niewiarowski S.;
RT "Albogagregin A and B. Structure and interaction with human
RT platelets.";
RL Thromb. Haemost. 79:609-613(1998).
CC -1- FUNCTION: Binds to platelet GPIb/IX receptor system and stimulates
CC agglutination.
CC -1- SUBUNIT: Heterodimer of the subunits 1, 2, 3 and 4,
CC disulfide-linked.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Contains 1 C-type lectin family domain.
DR HSSP; P23806; 11XX.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR003990; Pancreatis_ac.
DR Pfam; PF00059; lectin_c; 1.
DR PRINTS; PR01504; PNCREXITTSAP.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C-TYPE LECTIN_1; 1.
DR PROSITE; PS50041; C-TYPE LECTIN_2; 1.
KW Lectin.
FT DOMAIN 3 131 C-TYPE LECTIN (LONG FORM).
FT DISULFID 4 15 BY SIMILARITY.
FT DISULFID 32 129 BY SIMILARITY.
FT DISULFID 104 121 BY SIMILARITY.
SQ SEQUENCE 134 AA; 15556 MW; 5F9D71FC86DE2435 CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 1; Length 134;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 QOCSSSE 16
Db 79 QOCSSSE 84

RESULT 10
YHTE YEAST STANDARD; PRT; 148 AA.
AC P38839;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 17.0 kDa protein in YCK1-A09 intergenic region.
GN YHR136C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCB1_TaxID=4932;
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=94378003; PubMed=8091229;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
RA Du Z., Favell A., Fulton L., Gattung S., Geisel C., Kirsten J.,
RA Kucaba T., Miller L., Jier M., Johnston L., Langston Y.,
RA Lacroille P., Louis E.J., Macri C., Mardis E., Menezes S., Mouser L.,
RA Nhan M., Rifkin L., Riles L., St Peter H., Trevaakis E., Vaughan K.,
RA Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
RA Vaudin M.;
RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
RT VIII.";
RL Science 265:2077-2082(1994).
-----

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; U10398; AAB68421.1; -
DR PIR; S48980; S48980.
DR SGD; S0001178; SPL2.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR GO; GO:0009266; P:response to temperature; IGI.
KW Hypothetical protein.
SQ SEQUENCE 148 AA; 17036 MW; 67C02AA833F842E9 CRC64;

Query Match 14.6%; Score 6; DB 1; Length 148;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 17 DEDIFE 22
Db 135 DEDIFE 140

RESULT 11
ID B2 DAUCA STANDARD; PRT; 207 AA.

AC P3707;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE B2 protein.
OS Daucus carota (Carrot).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; campanulids; Apiales; Apiaceae; Daucus.
RN NCB1_TaxID=4039;
RX [1]
RA SEQUENCE FROM N.A.
RA Schrader S., Kaldenhoff R., Richter G.;
RL Submitted (MAR-1993) to the EMBL/GenBank/DBJ databases.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X72385; CA51078.1; -
DR PIR; S32124; S32124.
SQ SEQUENCE 207 AA; 23742 MW; BF66B257D2A3C221 CRC64;

Query Match 14.6%; Score 6; DB 1; Length 207;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 DIFERT 24
Db 200 DIFERT 205

RESULT 12
YIL3_YEAST STANDARD; PRT; 209 AA.
AC P40479;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Probable protein-tyrosine phosphatase YIL113W (EC 3.1.3.48).
YIL113W.

OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Schizosaccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCB1_TaxID=4932;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX PubMed=9169870;

RA Chillingworth T., Connor R., Devlin K., Gentles S., Hamlin N.,
RA Harris D.B., Horsnell T., Hunt S., Jagels K., Jones M., Lye G.,
RA Moule S., Odell C., Pearson D., Rajandream M.A., Rice P., Rowley N.,
RA Skelton J., Smith V., Walsh S., Whitehead S., Barrett B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome IX.";
RL Nature 387:84-87(1997).

CC -1- CATALYTIC ACTIVITY. Protein tyrosine phosphatase + H(2)O = protein
CC tyrosine + phosphate.
CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; Z38125; CA86267.1; -
DR PIR; S48459; S48459.
DR SGD; S0001375; YIL113W.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0017017; F:MAP kinase phosphatase activity; IDA.
DR GO; GO:0000196; P:MAPK cascade (cell wall biogenesis); IGI.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR00387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR SMART; SM00195; DSPC; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hypothetical protein; Hydrolase.
FT ACT SITE 140 140 BY SIMILARITY.
SQ SEQUENCE 209 AA; 23885 MW; 678371462A890A69 CRC64;

Query Match 14.6%; Score 6; DB 1; Length 209;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 DERSTD 8
Db 29 DERSTD 34

RESULT 13
YES4_SCHPO STANDARD; PRT; 264 AA.
ID YES4_SCHPO
AC O14171;
DT 15-JUL-1998 (Rel. 36, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein C4D7.04c in chromosome I.
GN SPAC4D7.04c.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCB1_TaxID=4896;
RX [1]
RA SEQUENCE FROM N.A.
RP STRAIN=972;
RC MEDLINE=21848401; PubMed=11859360;

RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodson K.,
 RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jaseles K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Gymnopoulos B.,
 RA Weljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Leinrich H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Medler H., Wambutt R., Purrelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Leilaure V., Motlier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.,
 RA "The genome sequence of Schizosaccharomyces pombe.";
 RT Nature 415:871-880(2002).
 CC -1- SIMILARITY: BELONGS TO THE UPP SYNTHETASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: 298602; CAB1276.1; -
 CC PIR: T38795; T38795.
 DR GeneDB_Spomb: SPAC4D7.04c; -
 DR InterPro: IPR001441; UPP_synth.
 DR Pfam: PF01255; UPP_synthetase; 1.
 DR ProDom: PD003461; UPP_synth; 1.
 DR TIGRFAMs: TIGR00055; upps; 1.
 DR PROSITE: PS01066; UPP_SYNTHETASE; 1.
 KW Hypothetical protein; Transferase.
 SQ SEQUENCE 264 AA; 30543 MW; 7DCC6094B114D29D CRC64;
 Query Match 14.6%; Score 6; DB 1; Length 264;
 Best Local Similarity 100.0%; Pred. No. 41;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 17 DEDIFE 22
 DB 193 DEDIFE 198
 RESULT 14
 MTB2_BACAM STANDARD; PRT; 265 AA.
 AC P18051;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Modification methylase BamHI (EC 2.1.1.113) (N-4 cytosine-
 DE specific methyltransferase BamHI) (M.BamHI).
 GN BAMHIIM.
 OS Bacillus amyloliquefaciens.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_Taxid=1390;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H;
 RX MEDLINE=90326538; PubMed=2374727;

RA Connaughton J.F., Kaloss W.D., Vanek P.G., Nardone G.A.,
 RA Chirifian J.G.;
 RT "The complete sequence of the Bacillus amyloliquefaciens proviral H2,
 RT BamHI methylase gene.";
 RL Nucleic Acids Res. 18:4002-4002(1990).
 CC -1- FUNCTION: THIS METHYLASE RECOGNIZES THE DOUBLE-STRANDED SEQUENCE
 CC GATCC, CAUSES SPECIFIC METHYLATION ON C-5 ON BOTH STRANDS, AND
 CC PROTECTS THE DNA FROM CLEAVAGE BY THE BAMHI ENDONUCLEASE.
 CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + DNA N(4)-methylcytosine = S-
 CC adenosyl-L-homocysteine + DNA N(4)-methylcytosine.
 CC -1- MISCELLANEOUS: THIS METHYLASE IS ENCODED ON THE PROVIRAL PHAGE H2.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X53032; CA37205.1; -
 CC PIR: S10316; S10316.
 DR HSSP: P1409; IBOO.
 DR REBASE: 186; M.BamHI.
 DR InterPro: IPR001091; CM4_Mettransf.
 DR InterPro: IPR002295; D21N6_methylase.
 DR InterPro: IPR002941; N6/N4_Mtase.
 DR InterPro: IPR000051; SAM_Bind.
 DR Pfam: PF01555; N6 N4 Mtase; 1.
 DR PRINTS: PR00506; D21N6MTFRASE.
 DR PRINTS: PR00508; S21N4MTFRASE.
 DR PROSITE: PS00093; N4_MTASE; 1.
 KW Transferase; Methyltransferase; Restriction system.
 SQ SEQUENCE 265 AA; 30981 MW; 895A51217A814820 CRC64;
 Query Match 14.6%; Score 6; DB 1; Length 265;
 Best Local Similarity 100.0%; Pred. No. 41;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 21 FEETRAQ 26
 DB 32 FEETRAQ 37
 RESULT 15
 YDDK_BACSU STANDARD; PRT; 266 AA.
 AC P66618;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein yddk.
 GN YDDK.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_Taxid=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunat F., Ogasaawa N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Berteiro M.G., Beesler P., Bolotin A., Borchert S.,
 RA Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Connerion I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,

RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaeser P., Goffeau A., Goldlighty E.J., Grandi G.,
RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kaashara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauei C., Medigue C.,
RA Medina N., Meliado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Portwolik S., Prescott A.M.,
RA Pressecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivoita C., Roche B., Roche M., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowaka A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weltzenegger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumschein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
subtilis".
RL Nature 390:249-256(1997).

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AB001488; BAA19337.1; -
DR EMBL; Z99106; CAB12307.1; -
DR PIR; D69776; D69776.
DR Subtilist; BG12118; YddK.
DR InterPro; IPR000157; TIR_domain.
DR SMART; SM00255; TIR; 1
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 266 AA; 31288 MW; E65A1RC3D2CE8CF CRC64;

Query Match 14.6%; Score 6; DB 1; Length 266;
Best Local Similarity 100.0%; Pred.No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 17 DEDIFE 22
Db 158 DEDIFE 163

Search completed: July 24, 2003, 12:22:06
Job time : 4.70132 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 12:16:28 ; Search time 12.6061 Seconds
(without alignments)
839.285 Million cell updates/sec

Title: US-09-991-681-28

Perfect score: 41
Sequence: 1 EDDERSTDSQQCSSEDEDI.....EETAQVSPRGKEXKQWRAR 41

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size: 0

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_prodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	1770	4 Q9ULH6	Q9ULH6 homo sapien
2	33	80.5	589	4 Q8N4Y4	Q8N4Y4 homo sapien
3	33	80.5	592	4 Q96CH9	Q96CH9 homo sapien
4	7	17.1	179	16 Q8XM35	Q8XM35 clostridium
5	7	17.1	200	13 Q93J93	Q93J93 gallus gall
6	7	17.1	221	10 Q9C6Z5	Q9C6Z5 arabidopsis
7	7	17.1	266	2 Q54360	Q54360 streptomyce
8	7	17.1	280	3 Q8X077	Q8X077 neurospora
9	7	17.1	318	3 Q9Y7P4	Q9Y7P4 schizosacch
10	7	17.1	340	4 Q8IXQ4	Q8IXQ4 homo sapien
11	7	17.1	351	4 Q9C0F9	Q9C0F9 homo sapien
12	7	17.1	402	5 Q45396	Q45396 caenorhabdi
13	7	17.1	435	11 Q8BNY8	Q8BNY8 mus musculu
14	7	17.1	794	10 Q940B8	Q940B8 arabidopsis
15	7	17.1	799	10 Q9LUS1	Q9LUS1 arabidopsis
16	7	17.1	898	11 Q08721	Q08721 ratus norv

17	7	17.1	1524	10 Q8RYN2	Q8RYN2 oryza sativ
18	7	17.1	1591	11 P97868	P97868 mus musculu
19	6	14.6	51	16 Q9KBL0	Q9KBL0 bacillus ha
20	6	14.6	57	4 Q9UNX2	Q9UNX2 homo sapien
21	6	14.6	65	17 Q8ZV17	Q8ZV17 pyrobaculum
22	6	14.6	68	11 Q922X7	Q922X7 mus musculu
23	6	14.6	69	17 Q8PUC7	Q8PUC7 methanobact
24	6	14.6	89	9 Q9B049	Q9B049 mycobacteri
25	6	14.6	96	15 Q36Z08	Q36Z08 human immun
26	6	14.6	111	16 Q8PAC7	Q8PAC7 xanthomonas
27	6	14.6	117	10 Q9LGA2	Q9LGA2 oryza sativ
28	6	14.6	120	16 Q8NMW0	Q8NMW0 staphylococ
29	6	14.6	123	2 Q52132	Q52132 escherichia
30	6	14.6	123	16 Q85638	Q85638 escherichia
31	6	14.6	124	12 Q65327	Q65327 oryza pseu
32	6	14.6	126	4 Q8N8D9	Q8N8D9 homo sapien
33	6	14.6	133	4 Q15908	Q15908 homo sapien
34	6	14.6	135	2 Q93P25	Q93P25 anaplasma m
35	6	14.6	135	5 Q8IQR5	Q8IQR5 drosophila
36	6	14.6	136	13 Q91841	Q91841 agkistrodon
37	6	14.6	148	3 P87290	P87290 saccharomyc
38	6	14.6	153	2 Q85178	Q85178 spirillum v
39	6	14.6	154	6 Q8MMW9	Q8MMW9 bos taurus
40	6	14.6	154	13 Q9YGN5	Q9YGN5 agkistrodon
41	6	14.6	155	16 Q32248	Q32248 bacillus su
42	6	14.6	158	16 Q9EYD0	Q9EYD0 escherichia
43	6	14.6	159	11 Q8CHY9	Q8CHY9 mus musculu
44	6	14.6	163	10 Q9ARN1	Q9ARN1 oryza sativ
45	6	14.6	165	2 P72251	P72251 rhodospirill

ALIGNMENTS

RESULT 1					
Q9ULH6	Q9ULH6	PRELIMINARY;	PRT;	1770 AA.	
AC	Q9ULH6; Q96P46;				
DT	01-MAY-2000 (TREMBlrel. 13, Created)				
DT	01-OCT-2002 (TREMBlrel. 22, Last sequence update)				
DT	01-OCT-2002 (TREMBlrel. 22, Last annotation update)				
DE	BiG3.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Hong W.;				
RT	"K1RA1244 as a novel distantly related member (BiG3) of the BiG1/Sec7p				
RL	subfamily of Arp GPFs.";				
DR	Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; AF413080; AL04174.1; -.				
DR	InterPro; IPR00904; Sec7.				
DR	SMART; SM00222; Sec7; 1.				
KW	Hypothetical protein.				
SO	SEQUENCE	1770 AA;	195845 MW;	58996R36AF92AB4 CRC64;	
Qy	Query Match	100.0%;	Score 41;	DB 4;	Length 1770;
Db	Best Local Similarity	100.0%;	Pred. No. 3.9e-36;		
	Matches 41;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
	1 EDDERSTDSQQCSSEDEDIFEETAQVSPRGKEXKQWRAR 41				
	1436 EDDERSTDSQQCSSEDEDIFEETAQVSPRGKEXKQWRAR 1476				
RESULT 2					
Q8N4Y4	Q8N4Y4	PRELIMINARY;	PRT;	589 AA.	
AC	Q8N4Y4;				
DT	01-OCT-2002 (TREMBlrel. 22, Created)				
DT	01-OCT-2002 (TREMBlrel. 22, Last sequence update)				

```

DE 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Similar to KIAA1244 protein (Fragment).
OC Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strausberg R.;
RT Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC033191; AAH33191.1; -.
FT NON TER
SQ SEQUENCE 589 AA; 66086 MW; 8041EA348DE65F7 CRC64;

Query Match
Best Local Similarity 100.0%; Score 33; DB 4; Length 589;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDERSTSSQCCSSEDEDFEETAGVSPRCK 33
DB 256 EDDERSTSSQCCSSEDEDFEETAGVSPRCK 288

RESULT 3
QY 096CH9 PRELIMINARY; PRT; 592 AA.
AC 096CH9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strausberg R.;
RT Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC014227; AAH14227.1; -.
KW Hypothetical protein.
FT NON TER
SQ SEQUENCE 592 AA; 66400 MW; F4A1E807B0DF47B5 CRC64;

Query Match
Best Local Similarity 100.0%; Score 33; DB 4; Length 592;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDERSTSSQCCSSEDEDFEETAGVSPRCK 33
DB 259 EDDERSTSSQCCSSEDEDFEETAGVSPRCK 291

RESULT 4
QY 08XM35 PRELIMINARY; PRT; 179 AA.
AC 08XM35;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Rubrychitin.
OS Rubrychitin.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=13 / Type A;
RX PubMed=11792842;
RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,

```

```

RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RT "Complete genome sequence of Clostridium perfringens, an anaerobic
RT flesh-eater."
RT Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
DR EMBL; AP003188; BAB80561.1; -.
DR InterPro; IPR003251; Rubrychitin.
DR InterPro; IPR006895; zf-Sec23_Sec24.
DR Pfam; PF02915; Rubrychitin; I.
DR Pfam; PF04810; zf-Sec23_Sec24; I.
DR ProDom; PD006198; Rubrychitin; I.
KW Complete proteome.
SQ SEQUENCE 179 AA; 20838 MW; 7D00B4A8B0D749FE CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 16; Length 179;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 DIFERTA 25
DB 43 DIFERTA 49

RESULT 5
QY 093393 PRELIMINARY; PRT; 200 AA.
AC 093393;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative calcium-activated potassium channel regulatory subunit
DE (Calcium-activated potassium channel beta subunit).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97224079; PubMed=9070660;
RA Oberst C., Weiskirchen R., Hartl M., Bister K.;
RT "Suppression in transformed avian fibroblasts of a gene (CO6) encoding
RT a membrane protein related to mammalian potassium channel regulatory
RT subunits."
RT Oncogene 14:1109-1116(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Oberst C., Bister K.;
RT "Chicken homolog of a quail gene (CO6) encoding a membrane protein
RT related to mammalian potassium channel regulatory subunits."
RT Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Bait S.L., Hudspeth A.J.;
RT Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF077369; AAC27490.1; -.
DR EMBL; AF420468; AAL16898.1; -.
DR InterPro; IPR003930; BK_channel_beta.
DR Pfam; PF03185; CASK; I.
KW Ionic channel.
SQ SEQUENCE 200 AA; 22663 MW; CFFD676158C5E0535 CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 13; Length 200;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 SEDEDIF 21
DB 73 SEDEDIF 79

RESULT 6
QY 09C6Z5

```

ID Q9C6Z5 PRELIMINARY; PRT; 221 AA.
AC Q9C6Z5;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Hypothetical 25.5 kDa protein.
GN T2H7.6.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RA White O., Ecker J.R., Palm C.J., Federpiel N.A., Kaul S., S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Creasy T.H., Dewar K.,
RA Dunn P., Etgu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huzar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
RA Miltchev J., Miranda M., Nguyen M., Nieman W.C., Osborne B.I.,
RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzer S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi L.J., Town C.D.,
RA Utecherack T., Van Aken S., Vaynsberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
thaliana";
RL Nature 408:816-820(2000).
RW EMBL; AC074176; AAG50860.1; -.
KW Hypothetical protein.
SQ SEQUENCE 221 AA; 25490 MW; 8647D307041EFP97 CRC64;

Query Match 17.1%; Score 7; DB 10; Length 221;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 29 PPRGKX 35
Db 4 PPRGKX 10

RESULT 7
ID Q54360 PRELIMINARY; PRT; 266 AA.
AC Q54360;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE lmbg protein.
GN lmbg.
OS Streptomyces lincolnensis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomyces.
OC NCBI_TaxID=1915;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=78-11;
RX MEDLINE=96020646; PubMed=8577749;
RA Peschke U., Schmidt H., Zhang H.Z., Piepersberg W.;
RT "Molecular characterization of the lincmycin-production gene cluster
of Streptomyces lincolnensis 78-11.";
RL Mol. Microbiol. 16:1137-1156(1995).
DR EMBL; X79146; CA55753.1; -.
SQ SEQUENCE 266 AA; 29268 MW; F21C81B6207DFB24 CRC64;

Query Match 17.1%; Score 7; DB 2; Length 266;

Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 24 TAQVSP 30
Db 11 TAQVSP 17

RESULT 8
ID Q8X0J7 PRELIMINARY; PRT; 280 AA.
AC Q8X0J7;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Hypothetical 30.7 kDa protein.
GN B11H24.060.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OC NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Algin V., Hohnselt J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RW EMBL; AL670005; CAD21276.1; -.
KW Hypothetical protein.
SQ SEQUENCE 280 AA; 30696 MW; 0C86CFA3FAA851 CRC64;

Query Match 17.1%; Score 7; DB 3; Length 280;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 13 CSSEDD 19
Db 258 CSSEDD 264

RESULT 9
ID Q9Y7P4 PRELIMINARY; PRT; 318 AA.
AC Q9Y7P4;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Mitochondrial ribosomal protein L2.
GN SPEC16C4.15.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OC NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972h-;
RA Purnelle B., Goffeau A., Wood V., Rajandream M.A., Barrell B.G.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL031335; CAB38631.1; -.
DR HSP; P04257; IRL2.
DR GeneDB_Spombe; SPCC16C4.15; -.
DR InterPro; IPR005880; L2_bact_org.
DR InterPro; IPR002171; Ribosomal L2.
DR Pfam; PF00181; Ribosomal L2; 1.
DR TIGRFAMs; TIGR01171; rplB_bact; 1.
DR PROSITE; PS00467; RIBOSOMAL_L2; 1.
KW Ribosomal protein.
SQ SEQUENCE 318 AA; 35211 MW; 4CBCT8529658F93E CRC64;

Query Match 17.1%; Score 7; DB 3; Length 318;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 30 PRGKEKR 36
 |||||
 DB 312 PRGKEKR 318

RESULT 10

O8IX04 PRELIMINARY; PRT; 340 AA.
 AC O8IX04:
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DE Similar to RIKEN CDNA 1200011118 gene.
 OS Homo sapiens (Human).
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OK NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=ovary;
 RA Strusberg R.;
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC039586; AAH39586.1; -.
 SQ SEQUENCE 340 AA; 38141 MW; F752A6A2917132E6 CRC64;

Query Match 17.1%; Score 7; DB 4; Length 340;
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 14 SSEDEDI 20
 |||||
 DB 140 SSEDEDI 146

RESULT 11

O9COF9 PRELIMINARY; PRT; 351 AA.
 AC O9COF9:
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DE Hypothetical protein KIAA1704 (Fragment).
 GN KIAA1704.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OK NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC MEDLINE=21082932; PubMed=11214970;
 RA Nagase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XIX.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro.";
 RL DNA Res. 7:347-355(2000).
 DR EMBL; AB051491; BAB21795.1; -.
 KW Hypothetical protein.
 FT NON TER 1
 SQ SEQUENCE 351 AA; 39406 MW; 6F5252CB5A6FE2F2 CRC64;

Query Match 17.1%; Score 7; DB 4; Length 351;
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 14 SSEDEDI 20
 |||||
 DB 151 SSEDEDI 157

RESULT 12

O45396 PRELIMINARY; PRT; 402 AA.
 AC O45396:
 DT 01-JUN-1998 (TREMBlrel. 06, Created)
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
 DE F22G12.2 protein.
 GN F22G12.2.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderiinae; Caenorhabditis.
 OK NCBI_TaxID=6239;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Leonard N.;
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
 RN (2)
 RP SEQUENCE FROM N.A.
 RC MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C. elegans: A platform for
 RT investigating biology.";
 RL Science 282:2012-2018(1998).
 DR EMBL; Z92831; CAB07366.1; -.
 DR WormPep; F22G12.2; CE15889.
 DR InterPro; IPR001584; Rve.
 DR Pfam; PF00665; rve; 1.
 SQ SEQUENCE 402 AA; 44636 MW; 22BAE51DAFA99F28 CRC64;

Query Match 17.1%; Score 7; DB 5; Length 402;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 18 EDIFEET 24
 |||||
 DB 188 EDIFEET 194

RESULT 13

O8BNY8 PRELIMINARY; PRT; 435 AA.
 AC O8BNY8:
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DE Retinoblastoma binding protein 6 (Fragment).
 GN Mus musculus (Mouse).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OK NCBI_TaxID=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Embryo;
 RA MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 DR EMBL; AK079129; BAC37553.1; -.
 FT NON TER 1
 SQ SEQUENCE 435 AA; 48396 MW; A0D10DB8C9AFCD4 CRC64;

Query Match 17.1%; Score 7; DB 11; Length 435;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 28 SPRGKE 34
 |||||
 DB 146 SPRGKE 152

```

1
RESULT 14
Q940B8 PRELIMINARY; PRT; 794 AA.
ID Q940B8
AC Q940B8
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Putative kinesin motor protein.
GN MGL6.9.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamada K., Banu J., Banno F., Dale J.M., Goldsmith A.D., Lee J.M.,
RA Onodera C.S., Quach H.L., Tang C., Toriumi M., Yamamura Y., Yu G.,
RA Yu S., Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C.,
RA Koesema E., Lam B., Lin J., Meyers M.C., Miranda M., Natusaka M.,
RA Nguyen M., Palm C.U., Sakurai T., Satou M., Seki M., Shimizu P.,
RA Southwick A., Tracy S.E., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RT "Full length cDNA of gene MGL6.9 (GI:11994617).";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY056129; AAL07208.1; -.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; kinesin; 1.
DR PRINTS; PRO0380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
KW ATP-binding; Coiled coil; Microtubules; Motor protein.
SQ SEQUENCE 794 AA; 89081 MW; 13CBd4DBEE2454B CRC64;

Query Match 17.1%; Score 7; DB 10; Length 794;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 VSPRCK 33
Db 643 VSPRCK 649

RESULT 15
Q9LUS1 PRELIMINARY; PRT; 799 AA.
ID Q9LUS1
AC Q9LUS1
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Similarity to kinesin protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Columbia;
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.,
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=Columbia;
RA MEDLINE=20277480; PubMed=10819329;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT features of the regions of 4,504,864 bp covered by sixty pl and TAC
RT clones.";
RL DNA Ref. 7:131-135(2000).

```

```

DR EMBL; AB022217; BAB02754.1; -.
DR HSSP; P17119; 3CAR.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; kinesin; 1.
DR PRINTS; PRO0380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
KW ATP-binding; Coiled coil; Microtubules; Motor protein.
SQ SEQUENCE 799 AA; 89720 MW; D700E00DB33A6647 CRC64;

Query Match 17.1%; Score 7; DB 10; Length 799;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 VSPRCK 33
Db 648 VSPRCK 654

Search completed: July 24, 2003, 12:25:47
Job time : 16.6061 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 11:44:42 ; Search time 13.3265 Seconds
(without alignments)
488.334 Million cell updates/sec

Title: US-09-991-681-28

Perfect score: 216
Sequence: 1 EDDERSTDSQQCSSEDEDI.....EETAQVSPPRGKRWRRAR 41

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
A_Geneseq_19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	216	100.0	41	23	AAO19166 Human prostate-spe
2	216	100.0	41	23	AAO19166 Human prostate-spe
3	216	100.0	518	19	AAO19166 Human prostate-spe
4	216	100.0	518	23	AAO19166 Human prostate-spe
5	216	100.0	518	22	AAO19166 Human prostate-spe
6	216	100.0	1807	22	ABG09728 Novel human diagno
7	216	100.0	1839	23	ABG09728 Novel human diagno
8	216	100.0	1982	22	ABG09731 Novel human diagno
9	69	31.9	192	20	AAO19166 Human prostate can

10	65.5	30.3	512	23	AAE19559
11	65.5	30.3	1585	22	ABG06075
12	65.5	30.3	1585	22	ABG10815
13	65.5	30.3	1598	22	ABG06618
14	65.5	30.3	1598	22	ABG09655
15	65.5	30.3	1647	22	ABG04001
16	65.5	30.3	1647	22	ABG04001
17	64	29.6	444	24	ABR41713
18	64	29.6	536	22	ABG92836
19	64	29.6	707	16	AAO19166
20	64	29.6	707	20	AAO19166
21	64	29.6	707	22	AAO19166
22	63	29.2	1418	23	ABG93257
23	62.5	28.9	642	22	ABG60513
24	62	28.7	268	22	ABG47541
25	62	28.7	268	22	ABG32527
26	62	28.7	268	22	ABG32527
27	62	28.7	268	22	ABG18177
28	62	28.7	268	22	AAO19166
29	62	28.7	268	22	AAO19166
30	62	28.7	268	22	AAO19166
31	62	28.7	268	22	AAO19166
32	62	28.7	268	22	AAO19166
33	62	28.7	268	22	AAO19166
34	62	28.7	268	22	AAO19166
35	60.5	28.0	512	15	AAO19166
36	60.5	28.0	512	15	AAO19166
37	60.5	28.0	512	15	AAO19166
38	60.5	28.0	512	15	AAO19166
39	59	27.3	635	23	ABG22101
40	59	27.3	957	22	AAO19166
41	59	27.3	1488	22	AAO19166
42	58.5	27.1	42	23	ABG31963
43	58.5	27.1	42	23	ABG31963
44	58.5	27.1	517	23	AAO19166
45	58.5	27.1	517	23	AAO19166

ALIGNMENTS

RESULT 1
AAO19166 standard; Protein: 41 AA.
ID AAO19166; standard; Protein: 41 AA.
AC AAO19166; standard; Protein: 41 AA.
XX 27-NOV-2002 (first entry)
DT 27-NOV-2002 (first entry)
XX Human prostate-specific PS118 protein fragment #2.
DE Human prostate-specific PS118 protein fragment #2.
XX Human prostate-specific PS118 protein fragment #2.
KW Human, prostate; prostate-specific sequence; prostate cancer; PS118;
KW EST; expressed sequence tag; cytosolic; gene therapy.
XX Homo sapiens.
OS Homo sapiens.
XX US2002086316-A1.
PN US2002086316-A1.
XX 04-JUL-2002.
XX 26-NOV-2001; 2001US-0991681.
XX 23-APR-1998; 98US-0065383.
XX 23-APR-1997; 97US-0842385.
XX (BIL) BILINGEL P. A.
XX (COHE) COHEN M.
XX (COLP) COLPITTS T. L.
XX (FRIE) FRIEDMAN P. N.
XX (GORD) GORDAN J.
XX (GRAN) GRANADOS E. N.
XX (HODG) HODGES S. C.
XX (KLAS) KLAS M. R.

PA (KRAT/) KRATOCHVIL J D.
 PA (ROBE/) ROBERTS-RAPP L.
 PA (RUSS/) RUSSELL J C.
 PA (STRO/) STROUPE S D.
 PI Billmel PA, Cohen M, Colpitts TL, Friedman PN, Gordan J,
 PI Granados EN, Hodges SC, Klass MR, Kratochvil JD, Roberts-Rapp L,
 PI Russell JC, Stroupe SD;
 XX WPI; 2002-665429/71.
 DR
 XX
 PT Novel PS18 polypeptide for detecting, diagnosing, staging, monitoring,
 PT prognosticating, preventing, treating, or determining predisposition of
 PT individual to diseases and conditions of prostate, e.g. prostate
 PT cancer -
 PS Claim 17; Page 43-44; 58bp; English.
 XX
 CC The present invention relates to a number of prostate-specific sequences
 CC derived from the human PS18 gene. These can be used in the detection,
 CC monitoring and treatment of prostate diseases, particularly prostate
 CC cancer. The present sequence is a PS18 protein fragment of the
 CC invention. The coding sequences of the invention were isolated from a
 CC prostate tissue expressed sequence tag (EST) library.
 XX
 SQ Sequence 41 AA;
 QY Query Match 100.0%; Score 216; DB 23; Length 41;
 Db Best Local Similarity 100.0%; Pred. No. 1e-21;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 EDDERSTDSQQCCSSEDEDFEETAOVSPRGKRRQWRAR 41
 1 EDDERSTDSQQCCSSEDEDFEETAOVSPRGKRRQWRAR 41
 RESULT 2
 ID AAM50810 standard; Protein; 41 AA.
 XX
 AC AAM50810;
 XX
 DT 01-MAY-2002 (first entry)
 XX
 DE PS18 prostate marker immunogenic polypeptide.
 XX
 KW PS18; prostate; marker; prostate cancer; tumour; metastasis;
 KW benign prostatic hyperplasia; prostatic intraepithelial neoplasia;
 KW prostatitis; human; diagnosis; therapy; vaccine; immunogen.
 XX
 OS Homo sapiens.
 XX
 PN US2001055758-A1.
 XX
 PD 27-DEC-2001.
 XX
 PF 23-APR-1998; 98US-0065383.
 XX
 PR 23-APR-1997; 97US-0842385.
 XX
 PA (BILL/) BILLING-MEDEL P A.
 PA (COHE/) COHEN M.
 PA (COPL/) COPLITTIS T L.
 PA (FRIE/) FRIEDMAN P N.
 PA (GORD/) GORDON J.
 PA (GRAN/) GRANADOS E N.
 PA (HODG/) HODGES S C.
 PA (KLAS/) KLAS M R.
 PA (KRAT/) KRATOCHVIL J D.
 PA (ROBE/) ROBERTS-RAPP L.
 PI Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN, Gordan J,
 PI Granados EN, Hodges SC, Klass MR, Kratochvil JD, Roberts-Rapp L,

PI Russell JC, Stroupe SD;
 XX
 DR WPI; 2002-187683/24.
 XX
 PT Detecting presence of target PS18 polynucleotide in test sample,
 PT useful for detecting, diagnosing, staging, monitoring, prognosticating,
 PT preventing or treating or determining predisposition to prostate
 PT disease -
 PS Claim 17; Page 43; 57bp; English.
 XX
 CC The present sequence is that of an immunogenic polypeptide
 CC comprising amino acids 184-224 of human prostate-specific PS18
 CC polypeptide (see AAM50809). A PS18 consensus sequence (see
 CC ABA91651) is found at least 12 times more often in prostate than
 CC in non-prostate tissue. PS18 polypeptides, polynucleotides,
 CC antibodies, agonists and inhibitors are useful for detecting,
 CC diagnosing, staging, monitoring, prognosticating, preventing and
 CC treating (including by genetic immunisation), or determining the
 CC predisposition of an individual to, diseases and conditions of the
 CC prostate, such as benign prostatic hyperplasia, prostatitis,
 CC prostatic intraepithelial neoplasia, prostate cancer, tumours and
 CC metastases. The PS18 polypeptides can be produced by expression
 CC of PS18 polynucleotides in transfected host cells, and
 CC immunogenic peptides are useful for raising PS18-specific
 CC antibodies of diagnostic use. The methods and reagents of the
 CC invention may provide an early means of detecting diseases of the
 CC prostate and may also provide new markers which can differentiate
 CC between the clinically important and unimportant prostate cancers
 CC without the use of surgery.
 XX
 SQ Sequence 41 AA;
 QY Query Match 100.0%; Score 216; DB 23; Length 41;
 Db Best Local Similarity 100.0%; Pred. No. 1e-21;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 EDDERSTDSQQCCSSEDEDFEETAOVSPRGKRRQWRAR 41
 1 EDDERSTDSQQCCSSEDEDFEETAOVSPRGKRRQWRAR 41
 RESULT 3
 ID AAW85472 standard; Protein; 518 AA.
 XX
 AC AAW85472;
 XX
 DT 25-MAR-2003 (updated)
 DT 25-FEB-1999 (first entry)
 XX
 DE PS18 protein encoded by consensus sequence.
 XX
 KW EST clone; PS18; prostate tumour tissue; prostatic disease; cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO9848054-A1.
 XX
 PD 29-OCT-1998.
 XX
 PF 23-APR-1998; 98WO-US08239.
 XX
 PR 23-APR-1997; 97US-0842385.
 XX
 PA (ABBO) ABBOTT LAB.
 PI Billing-medel PA, Cohen M, Colpitts TL, Friedman PN, Gordan J,
 PI Granados EN, Hodges SC, Klass MR, Kratochvil JD, Robertsrapp L,
 PI Russell JC, Stroupe SD;
 XX
 DR WPI; 1998-610006/51.
 DR N-PSDB; AAW82812.

XX New p18 nucleic acid and proteins - used for diagnosis and
PT treatment of prostatic disease, especially cancer, and also for drug
PT screening
XX
PS Claim 17; Page 93-94; 117pp; English.
XX
CC The present sequence is encoded by consensus pS18 sequence derived from
CC overlapping EST (expressed sequence tag) clones AAV82803-10. The clones
CC were identified from a cDNA library made from prostate tumour tissue.
CC Recombinant pS18 protein is used to detect pS18-specific antibodies,
CC to raise antibodies for detection of pS18 antigens, to screen for
CC specific binding agents (potential therapeutics), and to isolate specific
CC antibodies from serum. Detection of pS18 protein or nucleic acid, which
CC are prostate related, and altered or elevated in prostatic disease, is
CC used for detection, diagnosis, staging, monitoring and prognosis of
CC prostatic disease, particularly cancer, and to identify subjects at
CC risk.
CC (Updated on 25-MAR-2003 to correct PI field.)
CC
XX Sequence 518 AA;
SQ

Query Match 100.0%; Score 216; DB 19; Length 518;
Best Local Similarity 100.0%; Pred. No. 1.9e-20;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDERSTDSQCCSDEDEDIFEEETAQVSPRGKEXKQWRAR 41
Db 184 EDDERSTDSQCCSDEDEDIFEEETAQVSPRGKEXKQWRAR 224

RESULT 4
ID AAO19165 standard; Protein: 518 AA.
XX
AC AAO19165;
XX
DT 27-NOV-2002 (first entry)
XX
DE Human prostate-specific pS18 protein fragment #1.
XX
KW Human; prostate; prostate-specific sequence; prostate cancer; pS118;
KW EST; expressed sequence tag; cytostatic; gene therapy.
XX
OS Homo sapiens.
XX
PN US2002086316-A1.
XX
PD 04-JUL-2002.
XX
PF 26-NOV-2001; 2001US-0991681.
XX
PR 23-APR-1998; 98US-0065383.
PR 23-APR-1997; 97US-0842385.
XX
XX (BILL/) BILLINGEL P A.
PA (COHE/) COHEN M.
PA (COLP/) COLPITTS T L.
PA (FRIE/) FRIEDMAN P N.
PA (GORD/) GORDAN J.
PA (GRAN/) GRANADOS E N.
PA (HODG/) HODGES S C.
PA (KLAAS M R.
PA (KRAT/) KRATOCHVIL J D.
PA (ROBE/) ROBERTS-RAPP L.
PA (RUS/) RUSSELL J C.
PA (STRO/) STROUPE S D.
XX
XX Billingsel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;
PI Granados EN, Hodges SC, Klaas MR, Kratochvil JD, Roberts-Rapp L;
PI Russell JC, Stroupe SD;
XX WPI; 2002-665429/71.

XX Novel pS18 polypeptide for detecting, diagnosing, staging, monitoring,
PT prognosticating, preventing, treating, or determining predisposition of
PT individual to diseases and conditions of prostate, e.g. prostate
PT cancer
XX
PS Claim 17; Page 42-43; 58pp; English.
XX
CC The present invention relates to a number of prostate-specific sequences
CC derived from the human pS18 gene. These can be used in the detection,
CC monitoring and treatment of prostate diseases, particularly prostate
CC cancer. The present sequence is a pS18 protein fragment of the
CC invention. The coding sequences of the invention were isolated from a
CC prostate tissue expressed sequence tag (EST) library.
XX
SQ Sequence 518 AA;
SQ

Query Match 100.0%; Score 216; DB 23; Length 518;
Best Local Similarity 100.0%; Pred. No. 1.9e-20;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDERSTDSQCCSDEDEDIFEEETAQVSPRGKEXKQWRAR 41
Db 184 EDDERSTDSQCCSDEDEDIFEEETAQVSPRGKEXKQWRAR 224

RESULT 5
ID AAM50809 standard; Protein: 518 AA.
XX
AC AAM50809;
XX
DT 01-MAY-2002 (first entry)
XX
DE pS18 prostate marker partial sequence.
XX
KW pS18; prostate; marker; prostate cancer; tumour; metastasis;
KW benign prostatic hyperplasia; prostatic intraepithelial neoplasia;
KW prostaticitis; human; diagnosis; therapy; vaccine.
XX
OS Homo sapiens.
XX
PN US2001055758-A1.
XX
PD 27-DEC-2001.
XX
PF 23-APR-1998; 98US-0065383.
XX
PR 23-APR-1997; 97US-0842385.
XX
XX (BILL/) BILLING-MEDEL P A.
PA (COHE/) COHEN M.
PA (COLP/) COLPITTS T L.
PA (FRIE/) FRIEDMAN P N.
PA (GORD/) GORDAN J.
PA (GRAN/) GRANADOS E N.
PA (HODG/) HODGES S C.
PA (KLAAS M R.
PA (KRAT/) KRATOCHVIL J D.
PA (ROBE/) ROBERTS-RAPP L.
XX
XX Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;
PI Granados EN, Hodges SC, Klaas MR, Kratochvil JD, Roberts-Rapp L;
PI Russell JC, Stroupe SD;
XX WPI; 2002-187683/24.
DR N-PSDB; ABA91651.
XX
XX Detecting presence of target pS18 polynucleotide in test sample,
PT useful for detecting, diagnosing, staging, monitoring, prognosticating,
PT preventing or treating or determining predisposition to prostate
PT disease -
XX

PS Claim 17, Page 42-43; 57pp; English.

CC The present sequence is that of a human prostate-specific PS118
 CC polypeptide, as predicted from a partial consensus cDNA sequence
 CC (see ABA91651), and lacking the N-terminal region. The PS118
 CC consensus sequence is found at least 12 times more often in
 CC prostate than in non-prostate tissue. PS118 polypeptides,
 CC including derivatives of the present sequence, polynucleotides,
 CC antibodies, agonists and inhibitors are useful for detecting,
 CC diagnosing, staging, monitoring, prognosticating, preventing and
 CC treating (including by genetic immunisation), or determining the
 CC predisposition of an individual to, diseases and conditions of the
 CC prostate, such as benign prostatic hyperplasia, prostatitis,
 CC prostatic intraepithelial neoplasia, prostate cancer, tumours and
 CC metastases. The PS118 polypeptides can be produced by expression
 CC of PS118 polynucleotides in transfected host cells. The methods
 CC and reagents of the invention may provide an early means of
 CC detecting diseases of the prostate and may also provide new markers
 CC which can differentiate between the clinically important and
 CC unimportant prostate cancers without the use of surgery.

CC Sequence 518 AA;

SO Query Match 100.0%; Score 216; DB 23; Length 518;
 Best Local Similarity 100.0%; Pred. No. 1.9e-20;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDERSTDSQQCSSEDEDFEETAOVSPPRGKERKQWRAR 41
 DB 184 EDDERSTDSQQCSSEDEDFEETAOVSPPRGKERKQWRAR 224

RESULT 6
 ABOG9728
 ID ABOG9728 standard; Protein; 1807 AA.

XX AC ABOG9728;
 XX DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #9719.

KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KM food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.
 KM WO200175067-A2.
 XX PN 11-OCT-2001.
 XX PD 30-MAR-2001; 2001WO-US08631.
 XX PF 31-MAR-2000; 2000US-0540217.
 XX PR 23-AUG-2000; 2000US-0649167.
 XX XX (HYSE-) HYSEQ INC.
 XX PA Drmanac RT, Liu C, Tang YT;
 XX PI WPI; 2001-639362/73.
 XX DR N-PSDB; AAS73915.

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostic, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX PS Claim 20; SEQ ID No 40087; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (II) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABOG0010-ABG0377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 1807 AA;

SO Query Match 100.0%; Score 216; DB 22; Length 1807;
 Best Local Similarity 100.0%; Pred. No. 8e-20;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDERSTDSQQCSSEDEDFEETAOVSPPRGKERKQWRAR 41
 DB 1473 EDDERSTDSQQCSSEDEDFEETAOVSPPRGKERKQWRAR 1513

RESULT 7
 ABP64835
 ID ABP64835 standard; Protein; 1839 AA.

XX AC ABP64835;
 XX DT 25-FEB-2003 (first entry)

DE Human protein SEQ ID 495.

KM Human; expressed sequence tag; EST;
 KM haematopoietic disorder; central nervous system disease; viral infection;
 KM peripheral nervous system disease; non-healing wound; infectious disease;
 KM immune deficiency; immune disorder; bacterial infection; allergy; cancer;
 KM fungal infection; autoimmune disorder; coagulation disorder; noctropic;
 KM anti-allergic; anti-inflammatory; immunosuppressive; neuroprotective;
 KM cytostatic; haemostatic; virucide; antibacterial; fungicide;
 KM immunostimulant; cerebroprotective.

OS Homo sapiens.
 KM WO200259260-A2.
 XX PN 01-AUG-2002.
 XX PD 16-NOV-2001; 2001WO-US42950.
 XX PF 17-NOV-2000; 2000US-0714936.
 XX PR (HYSE-) HYSEQ INC.
 XX XX Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;
 XX PI Ren F, Xue AJ, Yang Y, Wehrman T, Drmanac RT;
 XX DR WPI; 2002-590824/63.
 XX DR N-PSDB; ABO99421.

PT New isolated polynucleotide, useful in research, diagnostic or
 PT therapeutic methods, e.g. preventing or treating disorders involving
 PT aberrant protein expression or biological activity -
 XX PS Claim 20; SEQ ID 495; 394pp; English.

CC The present invention relates to novel human coding sequences
 CC (ABQ99268-ABQ99608) and proteins (ABP64682-ABP65022). The sequences are
 CC useful in therapeutic, diagnostic and research methods. The
 CC polynucleotides may be used in the field of molecular biology as
 CC hybridisation probes, primers for PCR, for chromosome and gene mapping,
 CC for the recombinant production of protein, or in generation of anti-sense
 CC DNA or RNA. The polynucleotides are useful in diagnostics as expressed
 CC sequence tags (ESTs) for identifying expressed genes or for physical
 CC mapping of the human genome. The proteins may be used as molecular weight
 CC markers, or as nutritional sources or supplements. The proteins may be
 CC used to maintain and expand cell population in a totipotent or
 CC pluripotent state useful for re-engineering damaged or diseased
 CC tissues, transplantation, manufacture of bio-pharmaceuticals or the
 CC development of bio-sensors. The polynucleotides and proteins are useful
 CC for preventing, treating or ameliorating disorders involving aberrant
 CC protein expression or biological activity, e.g. hematopoietic disorders,
 CC central/peripheral nervous system diseases, mechanical and traumatic
 CC disorders, non-healing wounds, immune deficiencies and disorders,
 CC infectious diseases caused by viral, bacterial or fungal infection,
 CC autoimmune disorders, allergic reactions and conditions, coagulation
 CC disorders, or cancer. The polynucleotide sequences of the invention were
 CC assembled from ESTs isolated mainly by sequencing by hybridisation, and
 CC in some cases, sequences obtained from one or more public databases.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1839 AA;

Query Match 100.0%; Score 216; DB 23; Length 1839;
 Best Local Similarity 100.0%; Pred. No. 8.2e-20;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EDDERSTDSQCCSSEDEDFEETRAQVSPRGKXKRWRRAR 41
 |||||
 Db 1505 EDDERSTDSQCCSSEDEDFEETRAQVSPRGKXKRWRRAR 1545

RESULT 8
 ABG09731
 ID ABG09731 standard; Protein; 1982 AA.
 AC ABG09731;
 XX 13-FEB-2002 (first entry)
 DT
 XX Novel human diagnostic protein #9722.
 DE
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX Homo sapiens.
 OS
 XX MO200175067-A2.
 PN
 XX 11-OCT-2001.
 PD
 XX 30-MAR-2001; 2001WO-US08631.
 PF
 XX 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 XX (HYSE-) HYSEQ INC.
 PA
 PI Drmanec RT, Liu C, Tang YT;
 XX WPI; 2001-639362/73.
 DR N-PSDB; AAS73918.
 XX
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -

XX Claim 20; SEQ ID No 40090; 103pp; English.

PS The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG03077 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1982 AA;

Query Match 100.0%; Score 216; DB 22; Length 1982;
 Best Local Similarity 100.0%; Pred. No. 8.9e-20;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EDDERSTDSQCCSSEDEDFEETRAQVSPRGKXKRWRRAR 41
 |||||
 Db 1587 EDDERSTDSQCCSSEDEDFEETRAQVSPRGKXKRWRRAR 1627

RESULT 9
 AA48248
 ID AA48248 standard; Protein; 192 AA.
 AC AA48248;
 XX 08-DEC-1999 (first entry)
 DT
 XX Human prostate cancer-associated protein 34.
 DE
 XX Expressed sequence tag; EST; prostate tumor; antitumor; treatment;
 KW gene therapy; tissue specificity human.
 XX Homo sapiens.
 OS
 XX DE19811193-A1.
 PN
 XX 16-SEP-1999.
 PD
 XX 10-MAR-1998; 98DE-1011193.
 PF
 XX 10-MAR-1998; 98DE-1011193.
 PR
 XX (META-) METAGEN GES GENOMFORSCHUNG MBH.
 PA
 PI Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosenthal A;
 XX WPI; 1999-519628/44.
 DR N-PSDB; AAZ33451.
 XX
 XX New nucleic acid expressed at high level in prostatic tumor tissue and
 PT encoded polypeptides, useful for treating cancer and screening for
 PT therapeutic agents -
 PS Claim 22; 128; 166pp; German.
 XX This invention describes novel nucleic acid sequences (A) that are

XX

5' sequence 1585 AA/

Query Match 30.3%; Score 65.5; DB 22; Length 1585;
Best Local Similarity 35.4%; Pred. No. 11;
Matches 17; Conservative 6; Mismatches 18; Indels 7; Gaps 1;
Oy 1 EDDERSTSSSQCS-----SEDEDIFEEETAQVSPPGKEKQWRAR 41
Db 1248 ENDENSLSSSSDCSKNKDEI SEESDIEKTEVKREPELQTRREMEER 1295

RESULT 12
ABG10815
ID ABG10815 standard; Protein; 1585 AA.
XX
AC ABG10815;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #10806.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN W0200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS75002.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 41174; 103bp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1585 AA;
Query Match 30.3%; Score 65.5; DB 22; Length 1585;
Best Local Similarity 35.4%; Pred. No. 11;

Matches 17; Conservative 6; Mismatches 18; Indels 7; Gaps 1;
Oy 1 EDDERSTSSSQCS-----SEDEDIFEEETAQVSPPGKEKQWRAR 41
Db 1248 ENDENSLSSSSDCSKNKDEI SEESDIEKTEVKREPELQTRREMEER 1295

RESULT 13
ABG06618
ID ABG06618 standard; Protein; 1598 AA.
XX
AC ABG06618;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #6609.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN W0200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS70805.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 36977; 103bp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1598 AA;
Query Match 30.3%; Score 65.5; DB 22; Length 1598;
Best Local Similarity 35.4%; Pred. No. 12;
Matches 17; Conservative 6; Mismatches 18; Indels 7; Gaps 1;

^
Fri Jul 25 15:43:55 2003

us-09-991-681-28.rag

Page 9

Search completed: July 24, 2003, 11:53:37
Job time : 14.3265 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 11:53:53 ; Search time 3.66179 Seconds
(without alignment)
473.743 Million cell updates/sec

Title: US-09-991-681-28

Perfect score: 216
Sequence: 1 EDERSTDSQQCSSEDEDI.....EETAQVSPPRKGRQWRAR 41

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA: *
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep: *
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep: *
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep: *
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep: *
5: /cgn2_6/ptodata/1/1aa/ECTUS.COMB.pep: *
6: /cgn2_6/ptodata/1/1aa/backfile1.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	216	100.0	41	US-09-065-383-28	Sequence 28, Appl
2	216	100.0	518	US-09-065-383-27	Sequence 27, Appl
3	63	29.2	714	US-08-990-114-3	Sequence 3, Appl
4	63	29.2	714	US-09-241-333-3	Sequence 3, Appl
5	59	27.3	923	US-09-252-991A-28964	Sequence 28964, A
6	58.5	27.1	566	US-09-996-243-41	Sequence 41, Appl
7	57	26.4	667	US-09-303-064-55	Sequence 55, Appl
8	57	26.4	667	US-09-086-503-55	Sequence 55, Appl
9	56	25.9	421	US-08-132-649-6	Sequence 6, Appl
10	56	25.9	421	US-08-767-579-6	Sequence 6, Appl
11	56	25.9	522	US-09-233-160-23	Sequence 23, Appl
12	55.5	25.7	418	US-08-618-170-16	Sequence 16, Appl
13	55	25.5	180	US-08-630-915A-196	Sequence 196, Appl
14	55	25.5	411	US-08-741-134-6	Sequence 6, Appl
15	54	25.0	226	US-08-431-080-26	Sequence 26, Appl
16	54	25.0	226	US-08-938-534-26	Sequence 26, Appl
17	54	25.0	226	US-09-345-294-26	Sequence 26, Appl
18	54	25.0	1085	US-08-431-080-28	Sequence 28, Appl
19	54	25.0	1085	US-08-938-534-28	Sequence 28, Appl
20	54	25.0	1085	US-09-345-294-28	Sequence 28, Appl
21	54	25.0	2414	US-08-227-536-2	Sequence 2, Appl
22	54	25.0	2414	PCT-US95-04682-2	Sequence 2, Appl
23	53	24.5	441	US-09-254-776B-79	Sequence 79, Appl
24	52.5	24.3	193	US-08-679-765-3	Sequence 3, Appl
25	52.5	24.3	193	US-09-196-525-3	Sequence 3, Appl
26	52.5	24.3	193	US-09-318-317-3	Sequence 3, Appl
27	52	24.1	710	US-09-079-812E-2	Sequence 2, Appl

28	51.5	23.8	152	4	US-08-504-617-5	Sequence 5, Appl
29	51.5	23.8	611	3	US-09-370-807-2	Sequence 2, Appl
30	51.5	23.8	611	4	US-09-921-259-2	Sequence 2, Appl
31	51.5	23.8	826	4	US-09-894-988A-47	Sequence 47, Appl
32	51.5	23.8	2476	2	US-08-276-967-2	Sequence 2, Appl
33	51	23.6	262	4	US-09-134-001C-3237	Sequence 3237, Ap
34	51	23.6	466	4	US-09-252-991A-17543	Sequence 17543, A
35	51	23.6	630	4	US-09-328-352-4238	Sequence 4238, Ap
36	50.5	23.4	392	3	US-08-301-162-2	Sequence 2, Appl
37	50.5	23.4	392	4	US-09-461-240-2	Sequence 2, Appl
38	50.5	23.4	392	4	US-09-968-927-2	Sequence 2, Appl
39	50.5	23.4	426	1	US-08-615-170-6	Sequence 6, Appl
40	50.5	23.4	426	1	US-08-615-170-15	Sequence 15, Appl
41	50.5	23.4	428	3	US-08-301-162-18	Sequence 18, Appl
42	50.5	23.4	428	4	US-09-461-240-18	Sequence 18, Appl
43	50.5	23.4	428	4	US-09-968-927-18	Sequence 18, Appl
44	50.5	23.4	816	1	US-08-190-802A-54	Sequence 54, Appl
45	50.5	23.4	816	3	US-08-477-346-54	Sequence 54, Appl

ALIGNMENTS

RESULT 1
US-09-065-383-28
Sequence 28, Application US/09065383
Patent No. 6391543
GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GORDON, JULIAN
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KLAS, MICHAEL R.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: ROBERTS-RAPP, LISA
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/065,383
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/842,385
FILING DATE: 23-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6084, US, P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 41 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6391543e
US-09-065-383-28

Query Match 100.0%; Score 216; DB 4; Length 41;
Best Local Similarity 100.0%; Pred. No. 6.7e-22;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EDDERSTDSSQCCSSEDEDFEETAQVSPPRGKRRQWRAR 41
Db 1 EDDERSTDSSQCCSSEDEDFEETAQVSPPRGKRRQWRAR 41

RESULT 2
US-09-065-383-27
Sequence 27, Application US/09065383
Patent No. 6391543

GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA

APPLICANT: COHEN, MAURICE

APPLICANT: COLPITS, TRACEY L.

APPLICANT: FRIEDMAN, PAULA N.

APPLICANT: GORDON, JULIAN

APPLICANT: GRANADOS, EDWARD N.

APPLICANT: HODGES, STEVEN C.

APPLICANT: KASS, MICHAEL R.

APPLICANT: KRATOCHVIL, JON D.

APPLICANT: ROBERTS-RAPP, LISA

APPLICANT: RUSSELL, JOHN C.

APPLICANT: STROUPE, STEPHEN D.

TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
FOR DETECTING DISEASES OF THE PROSTATE

NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories

STREET: 100 Abbott Park Road

CITY: Abbott Park

STATE: IL

COUNTRY: USA

ZIP: 60064-3500

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/065,383

FILING DATE:
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/842,385

FILING DATE: 23-APR-1997

ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.

REGISTRATION NUMBER: 35,441

REFERENCE/DOCKET NUMBER: 6084.US.PI

TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729

TELEFAX: 847/938-2623

TELEX:

INFORMATION FOR SEQ ID NO: 27:

SEQUENCE CHARACTERISTICS:
LENGTH: 518 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: No. 6391543e

US-09-065-383-27

Query Match 100.0%; Score 216; DB 4; Length 518;
Best Local Similarity 100.0%; Pred. No. 1.2e-20;

Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EDDERSTDSSQCCSSEDEDFEETAQVSPPRGKRRQWRAR 41
Db 184 EDDERSTDSSQCCSSEDEDFEETAQVSPPRGKRRQWRAR 224

RESULT 3
US-08-990-114-3
Sequence 3, Application US/08990114
Patent No. 5932475

GENERAL INFORMATION:
APPLICANT: Bandman, Olga

APPLICANT: Yue, Henry

APPLICANT: Cortley, Neil C.

APPLICANT: Shah, Purvi

TITLE OF INVENTION: HUMAN NUCLEOLIN-LIKE PROTEIN

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/990,114

FILING DATE: Herewith

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0451 US

TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555

TELEFAX: 650-845-4166

TELEX:

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 714 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:
LIBRARY: GenBank

CLONE: 128842

US-08-990-114-3

Query Match 29.2%; Score 63; DB 2; Length 714;
Best Local Similarity 32.4%; Pred. No. 2.8;
Matches 12; Conservative 8; Mismatches 17; Indels 0; Gaps 0;

Qy 1 EDDERSTDSSQCCSSEDEDFEETAQVSPPRGKRRQ 37
Db 248 EDDERSTDSSQCCSSEDEDFEETAQVSPPRGKRRQ 264

RESULT 4
US-09-241-333-3
Sequence 3, Application US/09241333
Patent No. 6313266

GENERAL INFORMATION:
APPLICANT: Bandman, Olga

APPLICANT: Yue, Henry

APPLICANT: Cortley, Neil C.

APPLICANT: Shah, Purvi
TITLE OF INVENTION: HUMAN NUCLEOLIN-LIKE PROTEIN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESS: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/241,333
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/990,114
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0451 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 714 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 128842
US-09-241-333-3

Query Match 29.2%; Score 63; DB 4; Length 714;
Best Local Similarity 32.4%; Pred. No. 2.8;
Matches 12; Conservative 8; Mismatches 17; Indels 0; Gaps 0;

Qy 1 EDDERSTSSQCCSEDDIFETTAQVSPRGKEKQ 37
Db 248 EDEDEEEDEEEDEEEDEEEDEEPPVAPGKRKE 284

RESULT 5
US-09-252-991A-28964
Sequence 28964, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 28964
LENGTH: 923
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28964

Query Match 27.3%; Score 59; DB 4; Length 923;

Best Local Similarity 33.3%; Pred. No. 13;
Matches 11; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

Qy 4 ERSTDSSQCCSEDDIFETTAQVSPRGKEK 36
Db 795 QREQAQAQQAQGENKEQREASQSPSSGSSOR 827

RESULT 6
US-09-996-243-41
Sequence 41, Application US/09996243
Patent No. 6478825
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2730P1C3
CURRENT APPLICATION NUMBER: US/09/996,243
PRIOR FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025

PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19

PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Query Match 27.1%, Score 58.5; DB 4; Length 566;
Best Local Similarity 29.8%; Pred. No. 8.5;
Matches 14; Conservative 13; Mismatches 9; Indels 11; Gaps 2;

QY 1 EDDERS-----TDSQCCSEDEDFEETAQVSPPGKEXRQRR 41
DB 139 EDEHSGNDSDGSEPSERKTRLEEEIVEQTM-----RRQRREWEAR 180

RESULT 7

US-09-303-064-55
; Sequence 55, Application US/09303064
; Patent No. 6221619
; GENERAL INFORMATION:
; APPLICANT: MAINE, Gregory T.
; APPLICANT: HUNT, Jeffrey C.
; APPLICANT: BROJANAC, Susan
; APPLICANT: JYH-TSING SHEU, Michael
; APPLICANT: CHOYAN, Linda E.
; APPLICANT: TYNER, Joan D.
; APPLICANT: HOWARD, Lawrence V.
; APPLICANT: PARMLEY, Stephen F.
; APPLICANT: REMINGTON, Jack S.
; APPLICANT: ARAUJO, Fausto
; APPLICANT: SUZUKI, Yasuhiko
; APPLICANT: LI, Shuli
; TITLE OF INVENTION: ANTIGEN COCKTAILS, P35 AND USES THEREOF
; FILE REFERENCE: 6361.US.P1
; CURRENT APPLICATION NUMBER: US/09/303,064
; CURRENT FILING DATE: 1999-04-30
; EARLIER APPLICATION NUMBER: 09/086,503
; EARLIER FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 55
; LENGTH: 667
; TYPE: PRT
; ORGANISM: Toxoplasma gondii
US-09-303-064-55

Query Match 26.4%; Score 57; DB 3; Length 667;
Best Local Similarity 38.9%; Pred. No. 16;
Matches 14; Conservative 4; Mismatches 18; Indels 0; Gaps 0;

QY 1 EDDERSTDSQOCSSEDDIFETRAQVSPRGKXR 36
DB 265 EGEGETTESASSENEDDTFHDALQELPEGLLEVR 300

RESULT 8

US-09-086-503-55
; Sequence 55, Application US/09086503A
; Patent No. 6329157
; GENERAL INFORMATION:
; APPLICANT: MAINE, Gregory T.
; APPLICANT: HUNT, Jeffrey C.
; APPLICANT: BROJANAC, Susan
; APPLICANT: JYH-TSING SHEU, Michael
; APPLICANT: CHOYAN, Linda E.
; APPLICANT: TYNER, Joan D.
; APPLICANT: HOWARD, Lawrence V.
; TITLE OF INVENTION: ANTIGEN COCKTAILS AND USES THEREOF
; FILE REFERENCE: 6361.US.01
; CURRENT APPLICATION NUMBER: US/09/086,503A
; CURRENT FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 55
; LENGTH: 667
; TYPE: PRT
; ORGANISM: Toxoplasma gondii
US-09-086-503-55

Query Match 26.4%; Score 57; DB 4; Length 667;
Best Local Similarity 38.9%; Pred. No. 16;
Matches 14; Conservative 4; Mismatches 18; Indels 0; Gaps 0;

QY 1 EDDERSTDSQOCSSEDDIFETRAQVSPRGKXR 36
DB 265 EGEGETTESASSENEDDTFHDALQELPEGLLEVR 300

RESULT 9

US-08-132-649-6
; Sequence 6, Application US/08132649
; Patent No. 5585462
; GENERAL INFORMATION:
; APPLICANT: Londres, Constantine
; APPLICANT: Greenberg, Andrew S.
; APPLICANT: Kimmel, Alan R.
; APPLICANT: Egan, John J.
; TITLE OF INVENTION: CLONING OF PERILIPIN PROTEINS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Stewart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/132,649
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Garrett-Mackowski, Eugenia
; REGISTRATION NUMBER: 37,330
; REFERENCE/DOCKET NUMBER: 15280-145-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9500
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 421 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-132-649-6

Query Match 25.9%; Score 56; DB 1; Length 421;
Best Local Similarity 34.2%; Pred. No. 13;
Matches 13; Conservative 7; Mismatches 12; Indels 6; Gaps 1;

QY 1 EDDERSTDSQOCSSEDDI-----FEETRAQVSPRG 32
DB 298 EDHEDQDTGEDETEEELEETENKPSVVALPGPRG 335

RESULT 10

US-08-767-579-6
; Sequence 6, Application US/08767579
; Patent No. 6074842
; GENERAL INFORMATION:
; APPLICANT: Londres, Constantine
; APPLICANT: Greenberg, Andrew S.
; APPLICANT: Kimmel, Alan R.
; APPLICANT: Egan, John J.
; TITLE OF INVENTION: CLONING OF PERILIPIN PROTEINS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Stewart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/767,579
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Garrett-Mackowski, Eugenia
REGISTRATION NUMBER: 37,330
REFERENCE/DOCKET NUMBER: 15280-145-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
SEQUENCE FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 421 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-767-579-6

Query Match      25.9%; Score 56; DB 3; Length 421;
Best Local Similarity 34.2%; Pred. No. 13;
Matches 13; Conservative 7; Mismatches 12; Indels 6; Gaps 1;

QY      1 EDDERSTDSQCCSSEDEI-----FEETAQVSPRG 32
DB      298 EDHEDQDTGEGDTEEEBELTEENKFSEVALPGRG 335

RESULT 11
US-09-232-160-23
Sequence 23, Application US/09232160
GENERAL INFORMATION:
APPLICANT: Steve Daniel
APPLICANT: James Gilmore
APPLICANT: Susan G. Stuart
APPLICANT: Laura Stuve
TITLE OF INVENTION: DETECTION OF ALTERED EXPRESSION OF GENES REGULATING CELL
FILE REFERENCE: PA-0003 US
CURRENT APPLICATION NUMBER: US/09/232,160
CURRENT FILING DATE: 1999-01-15
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PERL Program
SEQ ID NO 23
LENGTH: 522
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE: -
OTHER INFORMATION: 3688209
US-09-232-160-23

Query Match      25.9%; Score 56; DB 4; Length 522;
Best Local Similarity 34.2%; Pred. No. 17;
Matches 13; Conservative 7; Mismatches 12; Indels 6; Gaps 1;

QY      1 EDDERSTDSQCCSSEDEI-----FEETAQVSPRG 32
DB      293 EDHEDQDTGEGDTEEEBELTEENKFSEVALPGRG 330

RESULT 12
US-08-615-170-16
Sequence 16, Application US/08615170
Patent No. 5776776
GENERAL INFORMATION:
APPLICANT: ORDAHL, Charles P.
APPLICANT: AZAKIE, Anthony
APPLICANT: MAR, Janet H.
APPLICANT: FARRANCE, Iain K.G.
APPLICANT: HALL, Deborah E.
```

```
APPLICANT: STEWART, Alexandre F.R.
APPLICANT: LARKIN, Sarah B.
TITLE OF INVENTION: DTEF-1 ISOFORMS AND USES THEREOF
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESS: Townsend and Townsend Kourie and Crew
STREET: Stewart Street Tower, One Market Plaza
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94105-1493
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/615,170
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/01526
FILING DATE: 06-FEB-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/191,493
FILING DATE: 04-FEB-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 2307U-053120
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 418 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-615-170-16

Query Match      25.7%; Score 55.5; DB 1; Length 418;
Best Local Similarity 31.8%; Pred. No. 15;
Matches 14; Conservative 6; Mismatches 15; Indels 9; Gaps 1;

QY      1 EDDERSTDSQCCSSEDEI-----EDIFETQVSPRGKX 35
DB      14 EDIERMSDSADKPIDNDAGVSPDIOSFOBALATYPGCRK 57

RESULT 13
US-08-630-915A-196
Sequence 196, Application US/08630915A
Patent No. 6309820
GENERAL INFORMATION:
APPLICANT: SPARKS, Andrew B.
APPLICANT: HOFEMAN, NO. 6309820H
APPLICANT: KAY, Brian K.
APPLICANT: FOWLER, Dana M.
APPLICANT: MCCONNELL, Stephen J.
TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
NUMBER OF SEQUENCES: 227
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
```

ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630.915A
FILING DATE: 03-APR-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Mirock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-174
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 196:
SEQUENCE CHARACTERISTICS:
LENGTH: 180 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-630-915A-196

Query Match 25.5%; Score 55; DB 4; Length 180;
Best Local Similarity 28.2%; Pred. No. 6.6; Indels 0; Gaps 0;
Matches 11; Conservative 10; Mismatches 18;

Qy 2 DDERSTDSQOCSEDEDFEETAOVSPPRGKEKQMPRA 40
Db 114 EEEGGSSESGSESDVAVETADGAEVKTQTPHNSA 152

RESULT 14
US-08-741-134-6
Sequence 6, Application US/08741134
Patent No. 5861498
GENERAL INFORMATION:
APPLICANT: Litwack, Gerald
APPLICANT: Alnemri, Emdad S.
APPLICANT: Fernandez-Alnemri, Teresa
TITLE OF INVENTION: IMMUNOPHILIN FKBP46 AND COMPOSITIONS FOR MAKING
TITLE OF INVENTION: AND
TITLE OF INVENTION: METHODS OF USING THE SAME
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5861498-its
STREET: One Liberty Place - 46th floor
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 3.11
SOFTWARE: WordPerfect for Windows 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/741.134
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/007.163
FILING DATE: 01-NOV-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Deluca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: T0U-2090
TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 411 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-741-134-6

Query Match 25.5%; Score 55; DB 2; Length 411;
Best Local Similarity 29.7%; Pred. No. 17;
Matches 11; Conservative 9; Mismatches 15; Indels 2; Gaps 1;

Qy 1 EDDERSTDSQOCSEDEDFEETAOVSPPRGKEKQ 37
Db 221 EDEEDNDGEEBOEEEEE--EKEEYKPPKSKKE 255

RESULT 15
US-08-431-080-26
Sequence 26, Application US/08431080
Patent No. 569886
GENERAL INFORMATION:
APPLICANT: Gottschling, Daniel E.
APPLICANT: Singer, William S.
TITLE OF INVENTION: Telomerase Compositions and Methods
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TEXAS
COUNTRY: UNITED STATES OF AMERICA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/431.080
FILING DATE: Concurrently Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SN 08/326,781
FILING DATE: October 20, 1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Parker, David L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: ARCD:155/PAR
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 226 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-431-080-26

Query Match 25.0%; Score 54; DB 1; Length 226;
Best Local Similarity 35.7%; Pred. No. 12;
Matches 15; Conservative 8; Mismatches 11; Indels 8; Gaps 2;

Qy 1 EDDERSTDSQOCSEDE--DIFETAOVSPPRGKEKQMPRA 40
Db 146 DDOGSDDSDSETSSDDENIDFVLTITQ-----RKGRAMKA 181

Fri Jul 25 15:43:57 2003

us-09-991-681-28.ra1

Page 8

Search completed: July 24, 2003, 12:16:22
Job time : 5.66179 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: July 24, 2003, 11:57:13 ; Search time 4.74231 Seconds
(without alignments)
1026.746 Million cell updates/sec

Title: US-09-991-681-28

Perfect score: 216
Sequence: 1 EDERSTRDSSQCCSEDEDI.....EETAQVSPPRGKRXQWRAR 41

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10C_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	63	29.2	714	9 US-09-978-242-3	Sequence 3, Appli
2	62.5	28.9	376	15 US-10-156-761-9888	Sequence 9888, Ap
3	62	28.7	268	9 US-09-864-761-33475	Sequence 33475, A
4	62	28.0	2799	14 US-10-151-736-4	Sequence 4, Appli
5	60.5	28.0	512	15 US-10-121-988-156	Sequence 156, App
6	59	27.3	957	9 US-09-801-574-80	Sequence 80, Appl
7	58.5	27.1	566	9 US-09-988-722-41	Sequence 41, Appl
8	58.5	27.1	566	9 US-09-988-723-41	Sequence 41, Appl
9	58.5	27.1	566	9 US-09-989-279-41	Sequence 41, Appl
10	58.5	27.1	566	9 US-09-989-727-41	Sequence 41, Appl
11	58.5	27.1	566	10 US-09-989-731-41	Sequence 41, Appl
12	58.5	27.1	566	10 US-09-989-732-41	Sequence 41, Appl
13	58.5	27.1	566	10 US-09-981-073-41	Sequence 41, Appl
14	58.5	27.1	566	10 US-09-990-442-41	Sequence 41, Appl
15	58.5	27.1	566	10 US-09-991-163-41	Sequence 41, Appl

16	58.5	27.1	566	10 US-09-993-604-41	Sequence 41, Appl
17	58.5	27.1	566	10 US-09-990-456-41	Sequence 41, Appl
18	58.5	27.1	566	10 US-09-989-721-41	Sequence 41, Appl
19	58.5	27.1	566	10 US-09-992-598-41	Sequence 41, Appl
20	58.5	27.1	566	10 US-09-989-293A-41	Sequence 41, Appl
21	58.5	27.1	566	10 US-09-989-735-41	Sequence 41, Appl
22	58.5	27.1	566	10 US-09-990-444-41	Sequence 41, Appl
23	58.5	27.1	566	10 US-09-991-181-41	Sequence 41, Appl
24	58.5	27.1	566	10 US-09-989-730-41	Sequence 41, Appl
25	58.5	27.1	566	10 US-09-990-436-41	Sequence 41, Appl
26	58.5	27.1	566	10 US-09-993-687-41	Sequence 41, Appl
27	58.5	27.1	566	11 US-09-989-734-41	Sequence 41, Appl
28	58.5	27.1	566	11 US-09-997-653-41	Sequence 41, Appl
29	58.5	27.1	566	11 US-09-993-667-41	Sequence 41, Appl
30	58.5	27.1	566	11 US-09-997-428-41	Sequence 41, Appl
31	58.5	27.1	566	11 US-09-997-666-41	Sequence 41, Appl
32	58.5	27.1	566	11 US-09-990-438-41	Sequence 41, Appl
33	58.5	27.1	566	11 US-09-990-562-41	Sequence 41, Appl
34	58.5	27.1	566	11 US-09-990-711-41	Sequence 41, Appl
35	58.5	27.1	566	11 US-09-989-726-41	Sequence 41, Appl
36	58.5	27.1	566	11 US-09-998-156-41	Sequence 41, Appl
37	58.5	27.1	566	11 US-09-990-437-41	Sequence 41, Appl
38	58.5	27.1	566	11 US-09-991-137-41	Sequence 41, Appl
39	58.5	27.1	566	11 US-09-997-514-41	Sequence 41, Appl
40	58.5	27.1	566	11 US-09-997-573-41	Sequence 41, Appl
41	58.5	27.1	566	11 US-09-991-172-41	Sequence 41, Appl
42	58.5	27.1	566	11 US-09-990-726-41	Sequence 41, Appl
43	58.5	27.1	566	11 US-09-997-559-41	Sequence 41, Appl
44	58.5	27.1	566	11 US-09-997-601-41	Sequence 41, Appl
45	58.5	27.1	566	11 US-09-990-443-41	Sequence 41, Appl

ALIGNMENTS

RESULT 1
US-09-978-242-3
Sequence 3, Application US/09978242
Patent No. US2002009856A1

GENERAL INFORMATION:

APPLICANT: Handman, Olga
Yue, Henry
Shah, Purvi
Corley, Neil C.
TITLE OF INVENTION: HUMAN NUCLEOLIN-LIKE PROTEIN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSER: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PasteSO for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/978, 242
FILING DATE: 15-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/241,333

FILING DATE: <Unknown>

APPLICATION NUMBER: 08/990,114

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0451 US

TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555


```

; APPLICANT: Sutherland, Lindfield
; APPLICANT: Watts, Colin K.
; TITLE OF INVENTION: No. US20020192160A1el Human Tumour Suppressor Gene
; FILE REFERENCE: RICE-0100CN
; CURRENT APPLICATION NUMBER: US/10/151,736
; PRIOR FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: 09/403,402
; PRIOR FILING DATE: 1999-10-21
; PRIOR APPLICATION NUMBER: PCT/AU98/00280
; PRIOR FILING DATE: 1998-04-20
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 2799
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-151-736-4

Query Match          28.7%; Score 62; DB 14; Length 2799;
Best Local Similarity 29.8%; Pred. No. 90;
Matches 17; Conservative 6; Mismatches 18; Indels 16; Gaps 2;

Cy 1 EDDERSTDS--SQCCSSEDEDIFETTAQVSP-----RGEKRGKRRAR 41
Db 1663 EDDSQSDSSDSSDSSSQSDIEDETFMLDEPLERTTSSHANGAQAAPRSMQWAVR 1719

RESULT 5
US-10-121-988-156
; Sequence 156, Application US/10121988
; Publication No. US20030068327A1
; GENERAL INFORMATION:
; APPLICANT: Hosken, Nancy Ann
; APPLICANT: McGowan, Patrick
; APPLICANT: Sleath, Paul R.
; APPLICANT: Moesman, Sally P.
; APPLICANT: Evans, Lawrence S.
; APPLICANT: Swanson, Ryan M.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF HERPES SIMPLEX VIRUS INFECTION
; FILE REFERENCE: 210121.538C1
; CURRENT APPLICATION NUMBER: US/10/121,988
; CURRENT FILING DATE: 2002-04-11
; NUMBER OF SEQ ID NOS: 183
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 156
; LENGTH: 512
; TYPE: PRT
; ORGANISM: HSV2
US-10-121-988-156

Query Match          28.0%; Score 60.5; DB 15; Length 512;
Best Local Similarity 34.1%; Pred. No. 23;
Matches 14; Conservative 7; Mismatches 13; Indels 7; Gaps 1;

Cy 3 DERSTDSQCCSSEDEDIFE-----ETAAVSPRGKEKR 36
Db 34 DPDESDSGECSSDDEDMEDPCGGAGAAIDAAIPKGPAPR 74

RESULT 6
US-09-801-574-80
; Sequence 80, Application US/09801574
; Patent No. US20020081592A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Peijing Jeremy
; APPLICANT: Page, David C.
; TITLE OF INVENTION: Reproduction-Specific Genes
; FILE REFERENCE: 0399.2007-002
; CURRENT APPLICATION NUMBER: US/09/801,574
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: 60/187,518
```

```

; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: 60/261,557
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 80
; LENGTH: 957
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-801-574-80

Query Match          27.3%; Score 59; DB 9; Length 957;
Best Local Similarity 31.6%; Pred. No. 67;
Matches 12; Conservative 7; Mismatches 19; Indels 0; Gaps 0;

Cy 1 EDDERSTDSQCCSSEDEDIFEETAAQVSPRGKEKRW 38
Db 836 ETDKKESSMLSKETEDLGEDETERAHSTLDEDERW 873

RESULT 7
US-09-989-722-41
; Sequence 41, Application US/09989722
; Patent No. US20020072067A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Baton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730PIC3
; CURRENT APPLICATION NUMBER: US/09/989,722
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
```

[illegible]

;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Query Match
Best Local Similarity 27.1%; Score 58.5; DB 9; Length 566;
Matches 14; Conservative 13; Mismatches 9; Indels 11; Gaps 2;

Qy 1 EDDERS-----TDSQCCSSEDEDIFETAAQVSPPRGKRWKRW 41
Db 139 EDEHSGNDSDGSESESEKRTLEETVEQTM-----RRQRREWEAR 180

RESULT 8
US-09-989-723-41
; Sequence 41, Application US/09989723
; Patent No. US20020072092A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Boetsen, David
; APPLICANT: Deenoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Geritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gutney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730P1C62
; CURRENT APPLICATION NUMBER: US/09/989,723
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02

;; PRIOR APPLICATION NUMBER: 60/087827
;; PRIOR FILING DATE: 1998-06-03
;; PRIOR APPLICATION NUMBER: 60/088021
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088025
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088026
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088028
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088029
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088030
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088033
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088326
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088167
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088202
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088212
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088217
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088655
;; PRIOR FILING DATE: 1998-06-09
;; PRIOR APPLICATION NUMBER: 60/088734
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088738
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088742
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088810
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088824
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088826
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088858
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088861
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088876
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/089105
;; PRIOR FILING DATE: 1998-06-12
;; PRIOR APPLICATION NUMBER: 60/089440
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089512
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089514
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089532
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089538
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089598
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089599
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089600
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089653
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089801
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089907
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089908
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089947

PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Query Match 27.1%; Score 58.5; DB 9; Length 566;
Best Local Similarity 29.8%; Pred. No. 45;
Matches 14; Conservative 13; Mismatches 9; Indels 11; Gaps 2;

QY 1 EDDERS-----TSSQCCSEEDIFETNAVSPPRKCRKRRAR 41
Db 139 EDEHSNGSDSGSEPSKRTLEBIEVOTW-----RRRORREWEAR 180

RESULT 9
US-09-989-279-41
Sequence 41, Application US/09989279
Patent No. US20020072496A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Deenoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gunney, Austin L.
APPLICANT: Kijavini, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2730P1C56
CURRENT APPLICATION NUMBER: US/09/989, 279
CURRENT FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026

PRIOR APPLICATION NUMBER: 60/090925	
PRIOR FILING DATE: 1998-06-22	
PRIOR APPLICATION NUMBER: 60/090255	
PRIOR FILING DATE: 1998-06-22	
PRIOR APPLICATION NUMBER: 60/090255	
PRIOR FILING DATE: 1998-06-22	
PRIOR APPLICATION NUMBER: 60/090345	
PRIOR FILING DATE: 1998-06-23	
PRIOR APPLICATION NUMBER: 60/090355	
PRIOR FILING DATE: 1998-06-23	
PRIOR APPLICATION NUMBER: 60/090422	
PRIOR FILING DATE: 1998-06-24	
PRIOR APPLICATION NUMBER: 60/090431	
PRIOR FILING DATE: 1998-06-24	
PRIOR APPLICATION NUMBER: 60/090435	
PRIOR FILING DATE: 1998-06-24	
PRIOR APPLICATION NUMBER: 60/090444	
PRIOR FILING DATE: 1998-06-24	
PRIOR APPLICATION NUMBER: 60/090445	
PRIOR FILING DATE: 1998-06-24	
PRIOR APPLICATION NUMBER: 60/090477	
PRIOR FILING DATE: 1998-06-24	
PRIOR APPLICATION NUMBER: 60/090533	
PRIOR FILING DATE: 1998-06-24	
PRIOR APPLICATION NUMBER: 60/090540	
PRIOR FILING DATE: 1998-06-24	
PRIOR APPLICATION NUMBER: 60/090544	
PRIOR FILING DATE: 1998-06-24	
PRIOR APPLICATION NUMBER: 60/090555	
PRIOR FILING DATE: 1998-06-24	
PRIOR APPLICATION NUMBER: 60/090676	
PRIOR FILING DATE: 1998-06-25	
PRIOR APPLICATION NUMBER: 60/090697	
PRIOR FILING DATE: 1998-06-25	
PRIOR APPLICATION NUMBER: 60/090699	
PRIOR FILING DATE: 1998-06-25	
PRIOR APPLICATION NUMBER: 60/090699	
PRIOR FILING DATE: 1998-06-25	
PRIOR APPLICATION NUMBER: 60/090699	
PRIOR FILING DATE: 1998-06-25	
PRIOR APPLICATION NUMBER: 60/090863	
PRIOR FILING DATE: 1998-06-26	
PRIOR APPLICATION NUMBER: 60/090863	
PRIOR FILING DATE: 1998-06-26	
PRIOR APPLICATION NUMBER: 60/091366	
PRIOR FILING DATE: 1998-07-01	
PRIOR APPLICATION NUMBER: 60/091478	
PRIOR FILING DATE: 1998-07-02	
PRIOR APPLICATION NUMBER: 60/091544	
PRIOR FILING DATE: 1998-07-01	
PRIOR APPLICATION NUMBER: 60/091513	
PRIOR FILING DATE: 1998-07-02	
PRIOR APPLICATION NUMBER: 60/091622	
PRIOR FILING DATE: 1998-07-02	
PRIOR APPLICATION NUMBER: 60/091633	
PRIOR FILING DATE: 1998-07-02	
PRIOR APPLICATION NUMBER: 60/091979	
PRIOR FILING DATE: 1998-07-07	
PRIOR APPLICATION NUMBER: 60/091987	
PRIOR FILING DATE: 1998-07-07	
PRIOR APPLICATION NUMBER: 60/092182	
PRIOR FILING DATE: 1998-07-09	

```
Query Match      27.1% Score 58.5; DB 9; Length 566;
Best Local Similarity 29.8%; Pred. No. 45;
Matches 14; Conservative 13; Mismatches 9; Indels 11; Gaps 2.
```

Qy 1 EDDERS-----TDSQCCSSEDEDIFEETAQVSPPRGKRRQRAR 41
||| | : : : : ||
Db 139 EDDEHSNDSDGSPEPSEKRTLRLEETVEQT-----RRRGRREWEAR 180

RESULT 10

US-09-989-727-41
Sequence 41, Application US/09989727
Patent No. US2002072497A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Borstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gertsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gueney, Austin L.
APPLICANT: Kjaev, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730P1C65
CURRENT APPLICATION NUMBER: US/09/989,727
CURRENT FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355

PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090662
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Query Match 27.1%; Score 58.5; DB 9; Length 566;
Best Local Similarity 29.8%; Pred. No. 45;
Matches 14; Conservative 13; Mismatches 9; Indels 11; Gaps 2;

QY 1 EDDRS-----TSSQCCSSEDEDFETAVSPRGKRRQWRAR 41
Db 139 EDEHSNDSDGSEPSERKTRLEETIVEQTM-----RRQRREWEAR 180

RESULT 11
US-09-989-731-41
Sequence 41, Application US/09989731
Patent No. US20020103125A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David

APPLICANT: Deenoyers, Luc
APPLICANT: Baton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gottlsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas P.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Collin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin

TITLE OR INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730P1C70
CURRENT FILING DATE: 2001-11-20
CURRENT APPLICATION NUMBER: US/09/989,731
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202

APPLICANT: Godowski, Paul J.
 APPLICANT: Grimaldi, J. Christopher
 APPLICANT: Gurney, Austin L.
 APPLICANT: Kljavin, Ivar J.
 APPLICANT: Napier, Mary A.
 APPLICANT: Pan, James
 APPLICANT: Paoni, Nicholas F.
 APPLICANT: Roy, Margaret Ann
 APPLICANT: Stewart, Timothy A.
 APPLICANT: Tumas, Daniel
 APPLICANT: Maranabe, Colin K.
 APPLICANT: Williams, P. Mickey
 APPLICANT: Wood, William I.
 APPLICANT: Zhang, Zemin
 TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 FILE REFERENCE: P2730P1C57
 CURRENT APPLICATION NUMBER: US/09/989,732
 CURRENT FILING DATE: 2001-11-19
 PRIOR APPLICATION NUMBER: 60/049787
 PRIOR FILING DATE: 1997-06-16
 PRIOR APPLICATION NUMBER: 60/062250
 PRIOR FILING DATE: 1997-10-17
 PRIOR APPLICATION NUMBER: 60/065186
 PRIOR FILING DATE: 1997-11-12
 PRIOR APPLICATION NUMBER: 60/065311
 PRIOR FILING DATE: 1997-11-13
 PRIOR APPLICATION NUMBER: 60/066770
 PRIOR FILING DATE: 1997-11-24
 PRIOR APPLICATION NUMBER: 60/075945
 PRIOR FILING DATE: 1998-02-25
 PRIOR APPLICATION NUMBER: 60/078910
 PRIOR FILING DATE: 1998-03-20
 PRIOR APPLICATION NUMBER: 60/083322
 PRIOR FILING DATE: 1998-04-28
 PRIOR APPLICATION NUMBER: 60/084600
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/087106
 PRIOR FILING DATE: 1998-05-28
 PRIOR APPLICATION NUMBER: 60/087607
 PRIOR FILING DATE: 1998-06-02
 PRIOR APPLICATION NUMBER: 60/087609
 PRIOR FILING DATE: 1998-06-02
 PRIOR APPLICATION NUMBER: 60/087759
 PRIOR FILING DATE: 1998-06-02
 PRIOR APPLICATION NUMBER: 60/087827
 PRIOR FILING DATE: 1998-06-03
 PRIOR APPLICATION NUMBER: 60/088021
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088025
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088026
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088028
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088029
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088030
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088033
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088326
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088167
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088202
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088212
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088217
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088655
 PRIOR FILING DATE: 1998-06-09

PRIOR APPLICATION NUMBER: 60/088734
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088738
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088742
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088810
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088824
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088826
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088858
 PRIOR FILING DATE: 1998-06-11
 PRIOR APPLICATION NUMBER: 60/088861
 PRIOR FILING DATE: 1998-06-11
 PRIOR APPLICATION NUMBER: 60/088876
 PRIOR FILING DATE: 1998-06-11
 PRIOR APPLICATION NUMBER: 60/089105
 PRIOR FILING DATE: 1998-06-12
 PRIOR APPLICATION NUMBER: 60/089440
 PRIOR FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: 60/089512
 PRIOR FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: 60/089514
 PRIOR FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: 60/089532
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089538
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089598
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089599
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089600
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089653
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089801
 PRIOR FILING DATE: 1998-06-18
 PRIOR APPLICATION NUMBER: 60/089907
 PRIOR FILING DATE: 1998-06-18
 PRIOR APPLICATION NUMBER: 60/089908
 PRIOR FILING DATE: 1998-06-18
 PRIOR APPLICATION NUMBER: 60/089947
 PRIOR FILING DATE: 1998-06-19
 PRIOR APPLICATION NUMBER: 60/089948
 PRIOR FILING DATE: 1998-06-19
 PRIOR APPLICATION NUMBER: 60/089952
 PRIOR FILING DATE: 1998-06-19
 PRIOR APPLICATION NUMBER: 60/090246
 PRIOR FILING DATE: 1998-06-22
 PRIOR APPLICATION NUMBER: 60/090252
 PRIOR FILING DATE: 1998-06-22
 PRIOR APPLICATION NUMBER: 60/090254
 PRIOR FILING DATE: 1998-06-22
 PRIOR APPLICATION NUMBER: 60/090349
 PRIOR FILING DATE: 1998-06-23
 PRIOR APPLICATION NUMBER: 60/090355
 PRIOR FILING DATE: 1998-06-23
 PRIOR APPLICATION NUMBER: 60/090429
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090431
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090435
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090444
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090445
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090472
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090535

```

PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Query Match
Best Local Similarity 29.8%; Pred. No. 45; DB 10; Length 566;
Matches 14; Conservative 13; Mismatches 9; Indels 11; Gaps 2;

Qy 1 EDDERS-----TDSQCCSEDEDFEETACVSPRGKEXRQMPAR 41
Db 139 EDEHSQNDSDGSEPEKXTRLEELVEQTM-----RRQRHWEAR 180

RESULT 13
US-09-991-073-41
Sequence 41, Application US/09991073
Patent No. US20020127576A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Geider, Hanspeter
APPLICANT: Gerltsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kijavini, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.

APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2730PIC15
CURRENT APPLICATION NUMBER: US/09/991.073
PRIOR FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
```

PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24

PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Query Match 27.1%; Score 58.5; DB 10; Length 566;
Best Local Similarity 29.8%; Pred. No. 45;
Matches 14; Conservative 13; Mismatches 9; Indels 11; Gaps 2;

Qy 1 EDDRS-----TDSQCGSDEDDIFETPAQVSPPGKRWRR 41
Db 139 EDEHSGNDSGDGSEPSKTRLEBIVEQTM-----RRORREWEAR 180

RESULT 14
US-09-990-442-41
Sequence 41, Application US/09990442
Patent No. US20020132252A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Deenoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kijavini, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin

;; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;; TITLE OF INVENTION: Acids Encoding the Same
;; FILE REFERENCE: P2730P1C8
;; CURRENT APPLICATION NUMBER: US/09/990,442
;; PRIOR FILING DATE: 2001-11-14
;; PRIOR APPLICATION NUMBER: 60/049787
;; PRIOR FILING DATE: 1997-06-16
;; PRIOR APPLICATION NUMBER: 60/062250
;; PRIOR FILING DATE: 1997-10-17
;; PRIOR APPLICATION NUMBER: 60/065186
;; PRIOR FILING DATE: 1997-11-12
;; PRIOR APPLICATION NUMBER: 60/065311
;; PRIOR FILING DATE: 1997-11-13
;; PRIOR APPLICATION NUMBER: 60/066770
;; PRIOR FILING DATE: 1997-11-24
;; PRIOR APPLICATION NUMBER: 60/075945
;; PRIOR FILING DATE: 1998-02-25
;; PRIOR APPLICATION NUMBER: 60/078910
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/083322
;; PRIOR FILING DATE: 1998-04-28
;; PRIOR APPLICATION NUMBER: 60/084600
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/087106
;; PRIOR FILING DATE: 1998-05-28
;; PRIOR APPLICATION NUMBER: 60/087607
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087609
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087759
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087827
;; PRIOR FILING DATE: 1998-06-03
;; PRIOR APPLICATION NUMBER: 60/088021
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088025
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088026
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088028
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088029
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088030
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088033
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088326
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088167
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088202
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088212
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088217
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088555
;; PRIOR FILING DATE: 1998-06-09
;; PRIOR APPLICATION NUMBER: 60/088734
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088738
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088742
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088810
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088824
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088826
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088858
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088861
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088876
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/089105
;; PRIOR FILING DATE: 1998-06-12
;; PRIOR APPLICATION NUMBER: 60/089440
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089512
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089514
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089532
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089538
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089598
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089599
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089600
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089653
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089801
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089907
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089908
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089947
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089948
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089952
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/090246
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090252
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090254
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090349
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090355
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090429
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090431
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090435
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090444
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090445
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090472
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090535
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090540
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090542
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090676
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090678
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090690
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090694

PRIOR FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: 60/089512
 PRIOR FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: 60/089514
 PRIOR FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: 60/089532
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089538
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089598
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089599
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089600
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089653
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089801
 PRIOR FILING DATE: 1998-06-18
 PRIOR APPLICATION NUMBER: 60/089907
 PRIOR FILING DATE: 1998-06-18
 PRIOR APPLICATION NUMBER: 60/089908
 PRIOR FILING DATE: 1998-06-18
 PRIOR APPLICATION NUMBER: 60/089947
 PRIOR FILING DATE: 1998-06-19
 PRIOR APPLICATION NUMBER: 60/089948
 PRIOR FILING DATE: 1998-06-19
 PRIOR APPLICATION NUMBER: 60/089952
 PRIOR FILING DATE: 1998-06-19
 PRIOR APPLICATION NUMBER: 60/090246
 PRIOR FILING DATE: 1998-06-22
 PRIOR APPLICATION NUMBER: 60/090252
 PRIOR FILING DATE: 1998-06-22
 PRIOR APPLICATION NUMBER: 60/090254
 PRIOR FILING DATE: 1998-06-22
 PRIOR APPLICATION NUMBER: 60/090349
 PRIOR FILING DATE: 1998-06-23
 PRIOR APPLICATION NUMBER: 60/090355
 PRIOR FILING DATE: 1998-06-23
 PRIOR APPLICATION NUMBER: 60/090429
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090431
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090435
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090444
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090445
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090472
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090535
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090540
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090542
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090557
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090676
 PRIOR FILING DATE: 1998-06-25
 PRIOR APPLICATION NUMBER: 60/090678
 PRIOR FILING DATE: 1998-06-25
 PRIOR APPLICATION NUMBER: 60/090690
 PRIOR FILING DATE: 1998-06-25
 PRIOR APPLICATION NUMBER: 60/090694
 PRIOR FILING DATE: 1998-06-25
 PRIOR APPLICATION NUMBER: 60/090695
 PRIOR FILING DATE: 1998-06-25
 PRIOR APPLICATION NUMBER: 60/090696
 PRIOR FILING DATE: 1998-06-25
 PRIOR APPLICATION NUMBER: 60/090862
 PRIOR FILING DATE: 1998-06-26

PRIOR APPLICATION NUMBER: 60/090863
 PRIOR FILING DATE: 1998-06-26
 PRIOR APPLICATION NUMBER: 60/091360
 PRIOR FILING DATE: 1998-07-01
 PRIOR APPLICATION NUMBER: 60/091478
 PRIOR FILING DATE: 1998-07-02
 PRIOR APPLICATION NUMBER: 60/091544
 PRIOR FILING DATE: 1998-07-01
 PRIOR APPLICATION NUMBER: 60/091519
 PRIOR FILING DATE: 1998-07-02
 PRIOR APPLICATION NUMBER: 60/091626
 PRIOR FILING DATE: 1998-07-02
 PRIOR APPLICATION NUMBER: 60/091633
 PRIOR FILING DATE: 1998-07-02
 PRIOR APPLICATION NUMBER: 60/091978
 PRIOR FILING DATE: 1998-07-07
 PRIOR APPLICATION NUMBER: 60/091982
 PRIOR FILING DATE: 1998-07-07
 PRIOR APPLICATION NUMBER: 60/092182
 PRIOR FILING DATE: 1998-07-09

Query Match 27.1%; Score 58.5; DB 10; Length 566;
 Best Local Similarity 29.8%; Pred. No. 45;
 Matches 14; Conservative 13; Mismatches 9; Indels 11; Gaps 2;

Qy 1 EDDERS-----TSSQCCSSEDDIFETIAQVSPPRGKRRQPAR 41
 Db 139 EDEHSGNDSGSESESEKTRLEIEIVQTM-----RRQRREWEAR 180

Search completed: July 24, 2003, 12:18:16
 Job time : 6.94231 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: July 24, 2003, 11:49:53 ; Search time 61.7701 Seconds

(without alignments)
577.701 Million cell updates/sec

Title: US-09-991-681-28

Perfect score: 216

Sequence: 1 EDDERSTDSQQCSSEDEDI.....EETAQVSPRGRKQRAR 41

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 5580241 seqs, 870357830 residues

Total number of hits satisfying chosen parameters: 5580241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US097A_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US099A_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
28: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
29: /cgn2_6/ptodata/1/paa/US103_COMB.pep.*
30: /cgn2_6/ptodata/1/paa/US104_COMB.pep.*
31: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	216	100.0	41	US-08-842-385-7	Sequence 7, Appl1
2	216	100.0	41	US-09-991-681-28	Sequence 28, Appl1

3	216	100.0	467	12	US-08-842-385-6	Sequence 6, Appl1
4	216	100.0	518	25	US-09-991-681-27	Sequence 27, Appl1
5	216	100.0	1770	1	PCT-US03-01943-44	Sequence 44, Appl1
6	216	100.0	1770	27	US-10-144-198-44	Sequence 44, Appl1
7	216	100.0	1807	1	PCT-US01-08631-40087	Sequence 40087, A
8	216	100.0	1839	1	PCT-US01-42950-495	Sequence 495, App
9	216	100.0	1839	30	US-10-416-993-495	Sequence 495, App
10	216	100.0	1872	1	PCT-US03-04508-32	Sequence 32, Appl1
11	216	100.0	1982	1	PCT-US01-06631-40090	Sequence 40090, A
12	216	100.0	2221	1	PCT-US03-01943-30	Sequence 30, Appl1
13	216	100.0	2221	27	US-10-144-198-30	Sequence 30, Appl1
14	69	31.9	192	20	US-09-623-791-87	Sequence 87, Appl1
15	69	31.9	192	20	US-09-623-791A-87	Sequence 87, Appl1
16	69	31.9	192	27	US-10-131-487A-87	Sequence 87, Appl1
17	65.5	30.3	462	26	US-10-072-012-770	Sequence 770, App
18	65.5	30.3	489	1	PCT-US02-38445-42	Sequence 42, Appl1
19	65.5	30.3	496	26	US-10-072-012-771	Sequence 771, App
20	65.5	30.3	496	31	US-60-389-987-2147	Sequence 2147, App
21	65.5	30.3	496	31	US-60-412-418-2147	Sequence 2147, App
22	65.5	30.3	521	26	US-10-072-012-769	Sequence 769, App
23	65.5	30.3	522	31	US-60-340-187-528	Sequence 528, App
24	65.5	30.3	533	22	US-09-758-472-9466	Sequence 9466, App
25	65.5	30.3	533	28	US-10-235-926-9466	Sequence 9466, App
26	65.5	30.3	1585	1	PCT-US01-08631-36434	Sequence 36434, A
27	65.5	30.3	1585	1	PCT-US01-08631-36977	Sequence 36977, A
28	65.5	30.3	1598	1	PCT-US01-08631-40014	Sequence 40014, A
29	65.5	30.3	1598	1	PCT-US01-08631-34360	Sequence 34360, A
30	65.5	30.3	1647	1	PCT-US02-03987-15575	Sequence 15575, A
31	65	30.1	606	26	US-10-032-585-7631	Sequence 7631, App
32	65	30.1	606	26	US-10-072-851-15575	Sequence 15575, A
33	65	30.1	606	31	US-60-259-128-5024	Sequence 5024, App
34	65	30.1	606	31	US-60-259-128-5024	Sequence 5024, App
35	65	30.1	606	31	US-60-314-050-7631	Sequence 7631, App
36	65	30.1	607	27	US-10-179-131-6595	Sequence 6595, App
37	64	29.6	536	20	US-09-629-469A-11383	Sequence 11383, A
38	64	29.6	687	27	US-10-104-047-2651	Sequence 2651, App
39	64	29.6	706	19	US-09-538-092-957	Sequence 957, App
40	64	29.6	706	28	US-10-219-051B-7546	Sequence 7546, App
41	64	29.6	706	28	US-10-219-051B-11767	Sequence 11767, A
42	64	29.6	707	1	PCT-US03-10240-19	Sequence 19, Appl1
43	64	29.6	707	17	US-09-593-302-22	Sequence 22, Appl1
44	64	29.6	707	22	US-09-791-537-27328	Sequence 27328, A
45	64	29.6	707	23	US-09-825-886-22	Sequence 22, Appl1

ALIGNMENTS

RESULT 1
US-08-842-385-7
Sequence 7, Application US/08842385

GENERAL INFORMATION:
APPLICANT: Russell, John
TITLE OF INVENTION: COLPITERS, Tracey
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL FOR DETECTING DISEASE OF THE PROSTATE
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESSES:
ADDRESS: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/842,385
FILING DATE:
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Porembski, Priscilla E
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 6084.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/937-6365
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 41 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: None
US-08-842-385-7

Query Match 100.0%; Score 216; DB 12; Length 41;
Best Local Similarity 100.0%; Pred. No. 2.2e-19;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDERSTDSQCCSSEDEDFEETAQVSPPRGKEKRWRR 41
DB 1 EDDERSTDSQCCSSEDEDFEETAQVSPPRGKEKRWRR 41

RESULT 2
US-09-991-681-28
Sequence 28, Application US/09991681
GENERAL INFORMATION:
APPLICANT: BILLING-MEDEI, PATRICIA
COHEN, MAURICE
COLPITTS, TRACEY L.
FRIEDMAN, PAULA N.
GORDON, JULIAN
GRANDOS, EDWARD N.
HODGES, STEVEN C.
KLASS, MICHAEL R.
KRATOCHVIL, JON D.
ROBERTS-RAPP, LISA
TITLE OF INVENTION: FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESS: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/991,681
FILING DATE: 26-Nov-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/065,383
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6084.US.PI
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 41 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: None
SEQUENCE DESCRIPTION: SEQ ID NO: 28:
US-09-991-681-28

Query Match 100.0%; Score 216; DB 25; Length 41;
Best Local Similarity 100.0%; Pred. No. 2.2e-19;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDERSTDSQCCSSEDEDFEETAQVSPPRGKEKRWRR 41
DB 1 EDDERSTDSQCCSSEDEDFEETAQVSPPRGKEKRWRR 41

RESULT 3
US-08-842-385-6
Sequence 6, Application US/08842385
GENERAL INFORMATION:
APPLICANT: Russell, John
COLPITTS, TRACEY
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
FOR DETECTING DISEASE OF THE PROSTATE
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/842,385
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Porembski, Priscilla E
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 6084.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/937-6365
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: None
US-08-842-385-6

Query Match 100.0%; Score 216; DB 12; Length 467;
Best Local Similarity 100.0%; Pred. No. 4.8e-18;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDERSTDSQCCSSEDEDFEETAQVSPPRGKEKRWRR 41
DB 133 EDDERSTDSQCCSSEDEDFEETAQVSPPRGKEKRWRR 173

RESULT 4
US-09-991-681-27
; Sequence 27, Application US/09991681
; GENERAL INFORMATION:
; APPLICANT: BILLING-MEDEL, PATRICIA
; COHEN, MAURICE
; COLPITTS, TRACEY L.
; FRIEDMAN, PAULA N.
; GORDON, JULIAN
; GRANADOS, EDWARD N.
; HODGES, STEVEN C.
; KLAS, MICHAEL R.
; KRATOCHVIL, JON D.
; ROBERTS-RAPP, LISA
; TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
; FOR DETECTING DISEASES OF THE PROSTATE
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/991,681
; FILING DATE: 26-Nov-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/065,383
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Becker, Cheryl L.
; REGISTRATION NUMBER: 35,441
; REFERENCE/DOCKET NUMBER: 6084.US.P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847/935-1729
; TELEFAX: 847/935-2623
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 518 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: None
; SEQUENCE DESCRIPTION: SEQ ID NO: 27:
US-09-991-681-27
Query Match 100.0%; Score 216; DB 25; Length 518;
Best Local Similarity 100.0%; Pred. No. 5,5e-18;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EDDERSTDSSQCCSSEDEDFEETAOVSPPRGKEKRWRR 41
DB 184 EDDERSTDSSQCCSSEDEDFEETAOVSPPRGKEKRWRR 224
RESULT 5
PCT-US03-01943-44
; Sequence 44, Application PC/TUS0301943
; GENERAL INFORMATION:
; APPLICANT: ORIGENE TECHNOLOGIES INC
; TITLE OF INVENTION: CANCER GENES
; FILE REFERENCE: 3U 9U 901 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/01943
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: US 10/054,935

; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: US 60/356,130
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 10/102,946
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: US 10/117,229
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: US 10/144,198
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: US 10/197,824
; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44
; LENGTH: 1770
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-01943-44
Query Match 100.0%; Score 216; DB 1; Length 1770;
Best Local Similarity 100.0%; Pred. No. 2.6e-17;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EDDERSTDSSQCCSSEDEDFEETAOVSPPRGKEKRWRR 41
DB 1436 EDDERSTDSSQCCSSEDEDFEETAOVSPPRGKEKRWRR 1476
RESULT 6
US-10-144-198-44
; Sequence 44, Application US/10144198
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies Inc
; TITLE OF INVENTION: Regulated Prostate Cance Genes
; FILE REFERENCE: 9U 105 R1
; CURRENT APPLICATION NUMBER: US/10/144,198
; CURRENT FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 44
; LENGTH: 1770
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-198-44
Query Match 100.0%; Score 216; DB 27; Length 1770;
Best Local Similarity 100.0%; Pred. No. 2.6e-17;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EDDERSTDSSQCCSSEDEDFEETAOVSPPRGKEKRWRR 41
DB 1436 EDDERSTDSSQCCSSEDEDFEETAOVSPPRGKEKRWRR 1476
RESULT 7
PCT-US01-08631-40087
; Sequence 40087, Application PC/TUS0108631
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-049
; CURRENT APPLICATION NUMBER: PCT/US01/08631
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 40087
; LENGTH: 1807
; TYPE: PRT
; ORGANISM: Homo sapiens

FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (48)..(62)
; OTHER INFORMATION: Fusion glycoprotein F0 domain identified by eMATRIX,
; OTHER INFORMATION: accession number PF00523D, p-value=7.188e-10, raw score of 11.39
; NAME/KEY: DOMAIN
; LOCATION: (941)..(950)
; OTHER INFORMATION: Helper component proteinase domain identified by Pfam,
; OTHER INFORMATION: accession name Peptidase_C6, E-value=0.0056, Pfam score of 7.9
PCT-US01-08631-40087

Query Match 100.0%; Score 216; DB 1; Length 1807;
Best Local Similarity 100.0%; Pred. No. 2.6e-17;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EDDERSTDSSQCCSSEDEDIFETTAQVSPPRGKEKRWRR 41
Db 1473 EDDERSTDSSQCCSSEDEDIFETTAQVSPPRGKEKRWRR 1513

RESULT 8
PCT-US01-42950-495
; Sequence 495, Application PC/TUS0142950
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-096
; CURRENT APPLICATION NUMBER: PCT/US01/42950
; CURRENT FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 09/714,936
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 682
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 495
; LENGTH: 1839
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-42950-495

Query Match 100.0%; Score 216; DB 1; Length 1839;
Best Local Similarity 100.0%; Pred. No. 2.7e-17;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EDDERSTDSSQCCSSEDEDIFETTAQVSPPRGKEKRWRR 41
Db 1505 EDDERSTDSSQCCSSEDEDIFETTAQVSPPRGKEKRWRR 1545

RESULT 9
US-10-416-993-495
; Sequence 495, Application US/10416993
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-096
; CURRENT APPLICATION NUMBER: US/10/416,993
; CURRENT FILING DATE: 2003-11-16
; PRIOR APPLICATION NUMBER: 09/714,936
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 682
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 495
; LENGTH: 1839
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-416-993-495

Query Match 100.0%; Score 216; DB 30; Length 1839;
Best Local Similarity 100.0%; Pred. No. 2.7e-17;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EDDERSTDSSQCCSSEDEDIFETTAQVSPPRGKEKRWRR 41

Db 1505 EDDERSTDSSQCCSSEDEDIFETTAQVSPPRGKEKRWRR 1545

RESULT 10
PCT-US03-04508-32
; Sequence 32, Application PC/TUS0304508
; GENERAL INFORMATION:
; APPLICANT: IDEC PHARMACEUTICALS
; APPLICANT: GATELY, DENNIS
; TITLE OF INVENTION: PROSTATE SPECIFIC GENES AND THE USE THEREOF IN DESIGN
; FILE REFERENCE: 037003/0301985
; CURRENT APPLICATION NUMBER: PCT/US03/04508
; CURRENT FILING DATE: 2003-02-19
; PRIOR APPLICATION NUMBER: 60/357,140
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: 60/396,082
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 60/386,759
; PRIOR FILING DATE: 2002-06-10
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 1872
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-04508-32

Query Match 100.0%; Score 216; DB 1; Length 1872;
Best Local Similarity 100.0%; Pred. No. 2.8e-17;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EDDERSTDSSQCCSSEDEDIFETTAQVSPPRGKEKRWRR 41
Db 1538 EDDERSTDSSQCCSSEDEDIFETTAQVSPPRGKEKRWRR 1578

RESULT 11
PCT-US01-08631-40090
; Sequence 40090, Application PC/TUS0108631
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-049
; CURRENT APPLICATION NUMBER: PCT/US01/08631
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 40090
; LENGTH: 1982
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (11)..(25)
; OTHER INFORMATION: Fusion glycoprotein F0 domain identified by eMATRIX,
; OTHER INFORMATION: accession number PF00523D, p-value=7.188e-10, raw score of 11.39
; NAME/KEY: DOMAIN
; LOCATION: (1065)..(1074)
; OTHER INFORMATION: Helper component proteinase domain identified by Pfam,
; OTHER INFORMATION: accession name Peptidase_C6, E-value=0.0056, Pfam score of 7.9
PCT-US01-08631-40090

Query Match 100.0%; Score 216; DB 1; Length 1982;
Best Local Similarity 100.0%; Pred. No. 3e-17;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EDDERSTDSSQCCSSEDEDIFETTAQVSPPRGKEKRWRR 41

Db 1587 EDDERSTDSSQCCSEDEDIFEEFTRQVSPPRGKEKRRAR 1627

RESULT 12

PCT-US03-01943-30
 ; Sequence 30, Application PC/TUS0101943
 ; GENERAL INFORMATION:
 ; APPLICANT: ORIGENE TECHNOLOGIES INC
 ; TITLE OF INVENTION: CANCER GENES
 ; FILE REFERENCE: 3U 9U 901 PCT
 ; CURRENT APPLICATION NUMBER: PCT/US03/01943
 ; PRIOR FILING DATE: 2002-01-25
 ; PRIOR APPLICATION NUMBER: US 10/054,935
 ; PRIOR FILING DATE: 2002-01-25
 ; PRIOR APPLICATION NUMBER: US 60/356,130
 ; PRIOR FILING DATE: 2002-02-14
 ; PRIOR APPLICATION NUMBER: US 10/102,946
 ; PRIOR FILING DATE: 2002-03-22
 ; PRIOR APPLICATION NUMBER: US 10/117,229
 ; PRIOR FILING DATE: 2002-04-08
 ; PRIOR APPLICATION NUMBER: US 10/144,198
 ; PRIOR FILING DATE: 2002-05-14
 ; PRIOR APPLICATION NUMBER: US 10/197,824
 ; PRIOR FILING DATE: 2002-07-19
 ; NUMBER OF SEQ ID NOS: 102
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 30
 ; LENGTH: 2221
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 PCT-US03-01943-30

Query Match 100.0%; Score 216; DB 1; Length 2221;
 Best Local Similarity 100.0%; Pred. No. 3,4e-17;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EDDERSTDSSQCCSEDEDIFEEFTRQVSPPRGKEKRRAR 41
 Db 1887 EDDERSTDSSQCCSEDEDIFEEFTRQVSPPRGKEKRRAR 1927

RESULT 13
 ; US-10-144-198-30
 ; Sequence 30, Application US/10144198
 ; GENERAL INFORMATION:
 ; APPLICANT: Origene Technologies Inc
 ; TITLE OF INVENTION: Regulated Prostate Cancer Genes
 ; FILE REFERENCE: 9U 105 R1
 ; CURRENT APPLICATION NUMBER: US/10/144,198
 ; CURRENT FILING DATE: 2002-05-14
 ; NUMBER OF SEQ ID NOS: 44
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 30
 ; LENGTH: 2221
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-144-198-30

Query Match 100.0%; Score 216; DB 27; Length 2221;
 Best Local Similarity 100.0%; Pred. No. 3,4e-17;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EDDERSTDSSQCCSEDEDIFEEFTRQVSPPRGKEKRRAR 41
 Db 1887 EDDERSTDSSQCCSEDEDIFEEFTRQVSPPRGKEKRRAR 1927

RESULT 14
 ; US-09-623-791-87
 ; Sequence 87, Application US/09623791
 ; GENERAL INFORMATION:
 ; APPLICANT: SPECHT, THOMAS
 ; APPLICANT: HINZMANN, BERND

; APPLICANT: SCHMITT, ARMIN
 ; APPLICANT: PILARSKY, CHRISTIAN
 ; APPLICANT: DAHL, EDGAR
 ; APPLICANT: ROSENTHAL, ANDRE
 ; TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM PROSTATE TUMOR TISSUE
 ; FILE REFERENCE: ALBRE 11
 ; CURRENT APPLICATION NUMBER: US/09/623,791
 ; CURRENT FILING DATE: 2001-07-10
 ; PRIOR APPLICATION NUMBER: PCT/DE99/00721
 ; PRIOR FILING DATE: 1999-03-09
 ; NUMBER OF SEQ ID NOS: 201
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 87
 ; LENGTH: 192
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-623-791-87

Query Match 31.9%; Score 69; DB 20; Length 192;
 Best Local Similarity 100.0%; Pred. No. 5.2;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EDDERSTDSSQCC 13
 Db 180 EDDERSTDSSQCC 192

RESULT 15
 ; US-09-623-791A-87
 ; Sequence 87, Application US/09623791A
 ; GENERAL INFORMATION:
 ; APPLICANT: SPECHT, THOMAS
 ; APPLICANT: HINZMANN, BERND
 ; APPLICANT: PILARSKY, CHRISTIAN
 ; APPLICANT: DAHL, EDGAR
 ; APPLICANT: ROSENTHAL, ANDRE
 ; TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM PROSTATE TUMOR TISSUE
 ; FILE REFERENCE: ALBRE 11
 ; CURRENT APPLICATION NUMBER: US/09/623,791A
 ; CURRENT FILING DATE: 2000-09-08
 ; PRIOR APPLICATION NUMBER: PCT/DE99/00721
 ; PRIOR FILING DATE: 1999-03-09
 ; NUMBER OF SEQ ID NOS: 201
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 87
 ; LENGTH: 192
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-623-791A-87

Query Match 31.9%; Score 69; DB 20; Length 192;
 Best Local Similarity 100.0%; Pred. No. 5.2;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EDDERSTDSSQCC 13
 Db 180 EDDERSTDSSQCC 192

Search completed: July 24, 2003, 12:15:12
 Job time : 62.7701 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 11:54:33 ; Search time 1.20059 Seconds
(without alignment)
140.227 Million cell updates/sec

Title: US-09-991-681-28

Perfect score: 216
Sequence: 1 EDERSTDSQOCSEDEDI.....ETRAQVSPRGKRWRRAR 41

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 41799 seqs, 4106219 residues

Total number of hits satisfying chosen parameters: 41799

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	216	100.0	1872	6 US-10-367-978-32	Sequence 32, Appl
2	60.5	28.0	512	1 PCT-US03-11231-156	Sequence 156, App
3	60.5	28.0	512	1 PCT-US03-11231-243	Sequence 243, Appl
4	58.5	27.1	517	6 US-10-451-901-16	Sequence 16, Appl
5	58.5	27.1	767	6 US-10-294-433-359	Sequence 359, Appl
6	53.5	24.8	538	7 US-60-478-196-3253	Sequence 3253, App
7	52	24.1	413	6 US-10-294-433-316	Sequence 316, App
8	52	24.1	739	6 US-10-353-856-1	Sequence 1, Appl1
9	51.5	23.8	825	1 PCT-US03-11231-161	Sequence 161, Appl
10	51.5	23.8	826	1 PCT-US03-11231-47	Sequence 47, Appl
11	51.5	23.6	352	6 US-10-273-573-10381	Sequence 10381, A
12	51	23.6	1170	7 US-60-478-196-3182	Sequence 3182, App
13	50	23.1	918	6 US-09-200-650E-1	Sequence 1, Appl1
14	49.5	22.9	45	6 US-10-273-573-8148	Sequence 8148, App
15	49.5	22.9	375	6 US-10-294-433-308	Sequence 308, App
16	49.5	22.9	775	5 US-09-820-843B-58	Sequence 58, Appl
17	49	22.7	281	1 PCT-US02-41612A-576	Sequence 576, Appl
18	49	22.7	363	6 US-10-275-558A-17	Sequence 17, Appl
19	49	22.7	839	1 PCT-US02-41612A-574	Sequence 308, App
20	49	22.7	1190	6 US-10-294-433-270	Sequence 270, App
21	49	22.7	1617	6 US-10-451-207-5	Sequence 5, Appl1
22	48	22.2	112	6 US-10-273-573-7902	Sequence 7902, Appl
23	48	22.2	250	6 US-10-273-573-7905	Sequence 2905, App
24	48	22.2	400	6 US-10-451-862-2	Sequence 2, Appl1
25	48	22.2	872	6 US-10-372-209-16	Sequence 16, Appl
26	48	22.2	1170	6 US-10-273-573-9652	Sequence 9652, Appl

27	47.5	22.0	1954	1 PCT-US03-19027-2	Sequence 2, Appl1
28	47	21.8	396	7 US-60-478-196-3236	Sequence 3236, App
29	47	21.8	547	1 PCT-US03-10753-42	Sequence 42, Appl
30	47	21.8	870	7 US-60-479-073-319	Sequence 319, Appl
31	47	21.8	1430	1 PCT-US03-06962-36	Sequence 36, Appl
32	46.5	21.5	202	7 US-60-478-196-3320	Sequence 3320, App
33	46.5	21.5	1469	7 US-60-479-073-335	Sequence 335, App
34	46	21.3	300	1 PCT-US02-18638A-176	Sequence 176, App
35	46	21.3	300	1 PCT-US02-18638A-174	Sequence 174, App
36	46	21.3	314	1 PCT-US02-18638A-172	Sequence 172, App
37	46	21.3	314	5 US-09-981-845-1	Sequence 1, Appl1
38	45.5	21.1	423	6 US-10-273-573-3690	Sequence 3690, App
39	45.5	21.1	1605	6 US-10-273-573-10656	Sequence 10656, App
40	45.5	21.1	1654	6 US-10-273-573-10655	Sequence 10655, A
41	45	20.8	146	1 PCT-US03-11231-7	Sequence 7, Appl1
42	45	20.8	257	6 US-10-273-573-5844	Sequence 5844, App
43	45	20.8	303	6 US-10-273-573-5845	Sequence 5845, App
44	45	20.8	468	6 US-10-273-573-5843	Sequence 5843, App
45	45	20.8	856	7 US-60-479-073-428	Sequence 428, App

ALIGNMENTS

```

RESULT 1
US-10-367-978-32
Sequence 32, Application US/10367978
GENERAL INFORMATION:
APPLICANT: GATELY, DENNIS
TITLE OF INVENTION: PROSTATE SPECIFIC GENES AND THE USE THEREOF IN DESIGN
FILE REFERENCE: 037003-0301988
CURRENT APPLICATION NUMBER: US/10/367,978
PRIOR FILING DATE: 2003-02-19
PRIOR APPLICATION NUMBER: 60/357,140
PRIOR FILING DATE: 2002-02-19
PRIOR APPLICATION NUMBER: 60/396,082
PRIOR FILING DATE: 2002-07-17
PRIOR APPLICATION NUMBER: 60/386,759
PRIOR FILING DATE: 2002-06-10
NUMBER OF SEQ ID NOS: 89
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 32
LENGTH: 1872
TYPE: PRT
ORGANISM: Homo sapiens
US-10-367-978-32

Query Match          100.0%; Score 216; DB 6; Length 1872;
Best Local Similarity 100.0%; Pred. No. 1.1e-21;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1 EDERSTDSQOCSEDEDIFFETRAQVSPRGKRWRRAR 41
Db      1538 EDERSTDSQOCSEDEDIFFETRAQVSPRGKRWRRAR 1578

RESULT 2
PCT-US03-11231-156
Sequence 156, Application PC/TUS0311231
GENERAL INFORMATION:
APPLICANT: Corixa Corporation
APPLICANT: Day, Craig H.
APPLICANT: Hoehen, Nancy A.
APPLICANT: Parsons, Joseph M.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
FILE REFERENCE: 210121.53801PC
CURRENT APPLICATION NUMBER: PCT/US03/11231
CURRENT FILING DATE: 2003-04-09
NUMBER OF SEQ ID NOS: 267
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 156

```

Query Match	27.1%	Score 58.5	DB 6	Length 767
Best Local Similarity	34.2%	Pred. No. 1.2		
Matches 13	Conservative 9	Mismatches 15	Indels 1	Gaps 1
QY	1	EDDERSTDSQQCSF--DEDIFEEETAQVSPRGKEXRQ	37	

Db 508 EDDASTSTASGEEDNIDEDMEEDRAGLGPPLSHRPQ 545

RESULT 6
US-60-478-196-3253

; Sequence 3253, Application US/60478196

; GENERAL INFORMATION:

; APPLICANT: Jiang, Bo

; APPLICANT: Lemieux, Sebastien

; APPLICANT: Hu, Wengqi

; APPLICANT: Rosemer, Terry

; TITLE OF INVENTION: IDENTIFICATION OF ESSENTIAL GENES OF ASPERGILLUS FUNIGATUS AND ME

; FILE REFERENCE: 10182-026-888

; CURRENT APPLICATION NUMBER: US/60/478.196

; NUMBER OF SEQ ID NOS: 4000

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 3253

; LENGTH: 538

; TYPE: PRT

; ORGANISM: Aspergillus fumigatus

US-60-478-196-3253

Query Match 24.8%; Score 53.5; DB 7; Length 538;

Best Local Similarity 30.6%; Pred. No. 3.9;

Matches 15; Conservative 6; Mismatches 19; Indels 9; Gaps 1;

Oy 1 EDDERSTDSQCCSSEDEDI-----FEETAQVSPPRKEXKRWMA 40

Db 71 EDDERISLEEDLQSEDEDMQVSEVAGDARTLAVDTQAPSSKRRKA 119

RESULT 7

US-10-294-433-316

; Sequence 316, Application US/10294433

; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc

; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

; FILE REFERENCE: 792CIP4

; CURRENT APPLICATION NUMBER: US/10/294.433

; PRIOR FILING DATE: 2002-11-13

; PRIOR APPLICATION NUMBER: PCT/US01/14826

; PRIOR FILING DATE: 2001-05-16

; PRIOR APPLICATION NUMBER: 09/989,600

; PRIOR FILING DATE: 2001-11-21

; PRIOR APPLICATION NUMBER: 09/577,408

; PRIOR FILING DATE: 2000-05-18

; PRIOR APPLICATION NUMBER: 10/115,831

; PRIOR FILING DATE: 2002-04-02

; PRIOR APPLICATION NUMBER: 09/677,298

; PRIOR FILING DATE: 2000-09-22

; PRIOR APPLICATION NUMBER: 09/695,781

; PRIOR FILING DATE: 2000-10-24

; PRIOR APPLICATION NUMBER: 10/150,802

; PRIOR FILING DATE: 2002-05-15

; PRIOR APPLICATION NUMBER: 09/715,869

; PRIOR FILING DATE: 2000-11-17

; PRIOR APPLICATION NUMBER: 10/167,379

; PRIOR FILING DATE: 2002-06-10

; PRIOR APPLICATION NUMBER: 09/775,330

; PRIOR FILING DATE: 2001-02-01

; NUMBER OF SEQ ID NOS: 864

; SOFTWARE: Custom

; SEQ ID NO 316

; LENGTH: 413

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-294-433-316

Query Match 24.1%; Score 52; DB 6; Length 413;

Best Local Similarity 32.3%; Pred. No. 4.5;

Matches 10; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

Oy 1 EDDERSTDSQCCSSEDEDIFFETAAVSPPR 31

Db 365 EDSERISIDMEDDDDDDDLEDSISLSPRK 395

RESULT 8

US-10-353-856-1

; Sequence 1, Application US/10353856

; GENERAL INFORMATION:

; APPLICANT: Kawasaka, Yoshihiro

; APPLICANT: Jasenosky, Luke D.

; APPLICANT: Neumann, Gabriele

; APPLICANT: Wisconsin Alumni Research Foundation

; TITLE OF INVENTION: Filovirus Vectors and Noninfectious Filovirus-Based Particles

; FILE REFERENCE: 800.032US1

; CURRENT APPLICATION NUMBER: US/10/353.856

; PRIOR FILING DATE: 2003-01-29

; PRIOR APPLICATION NUMBER: US 60/353,972

; PRIOR FILING DATE: 2002-01-31

; NUMBER OF SEQ ID NOS: 49

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 739

; TYPE: PRT

; ORGANISM: Reston Ebola virus

US-10-353-856-1

Query Match 24.1%; Score 52; DB 6; Length 739;

Best Local Similarity 37.1%; Pred. No. 9;

Matches 13; Conservative 10; Mismatches 8; Indels 4; Gaps 2;

Oy 1 EDDERSTDSQCCSSEDEDIFFETAAVSPPRKEX 35

Db 525 KDDRRASDNNQ--SADSE--EQGYRRHRGPER 555

RESULT 9

PCT-US03-11231-161

; Sequence 161, Application PC/TUS0311231

; GENERAL INFORMATION:

; APPLICANT: Corixa Corporation

; APPLICANT: Day, Craig H.

; APPLICANT: Hosken, Nancy A.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND

; FILE REFERENCE: 210121.53801PC

; CURRENT APPLICATION NUMBER: PCT/US03/11231

; CURRENT FILING DATE: 2003-04-09

; NUMBER OF SEQ ID NOS: 267

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 161

; LENGTH: 825

; TYPE: PRT

; ORGANISM: HSV2

PCT-US03-11231-161

Query Match 23.8%; Score 51.5; DB 1; Length 825;

Best Local Similarity 33.3%; Pred. No. 12;

Matches 13; Conservative 6; Mismatches 17; Indels 3; Gaps 1;

Oy 2 DDERSTDSQCCSSEDEDIFFETAAVSPPRKEX 37

Db 57 DDDLHRDSTSEAGSTDTMEFAGLMDATPPARPARQ 95

RESULT 10

PCT-US03-11231-47

; Sequence 47, Application PC/TUS0311231

; GENERAL INFORMATION:

; APPLICANT: Corixa Corporation

```
; APPLICANT: Day, Craig H.
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Parsons, Joseph M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF HERPES SIMPLEX VIRUS INFECTION
; FILE REFERENCE: 210121.53801PC
; CURRENT APPLICATION NUMBER: PCT/US03/11231
; CURRENT FILING DATE: 2003-04-09
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 826
; TYPE: PRT
; ORGANISM: HSV-2
; PCT-US03-11231-47

Query Match      23.8%; Score 51.5; DB 1; Length 826;
Best Local Similarity 33.3%; Pred. No. 12;
Matches 13; Conservative 6; Mismatches 17; Indels 3; Gaps 1;

QY      2 DDERSTDSQQCSSEDEDIFEETAAQVSPPRGKEKQ 37
DB      57 DDLHRDSTSEAGSTDTETMEFAGLMDAATPPAPPAERQ 95

RESULT 11
US-10-273-573-10381
; Sequence 10381, Application US/10273573
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: US/10/273,573
; CURRENT FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 10381
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(352)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-273-573-10381

Query Match      23.6%; Score 51; DB 6; Length 352;
Best Local Similarity 39.4%; Pred. No. 5.2;
Matches 13; Conservative 5; Mismatches 7; Indels 8; Gaps 2;

QY      2 DDERSTDSQQCSSEDEDIFEETAAQVSPPRGKE 34
DB      143 EDEKETD-----EDDERFGEALQ-SPASGKQ 167

RESULT 12
US-60-478-196-3182
; Sequence 3182, Application US/60478196
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Lemieux, Sebastien
; APPLICANT: Hu, Wengli
; APPLICANT: Roemer, Terry
; TITLE OF INVENTION: IDENTIFICATION OF ESSENTIAL GENES OF ASPERGILLUS FUMIGATUS AND ME
; TITLE OF INVENTION: USE
; FILE REFERENCE: 10182-026-888
; CURRENT APPLICATION NUMBER: US/60/478,196
; CURRENT FILING DATE: 2003-06-13
; NUMBER OF SEQ ID NOS: 4000
```

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3182
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-60-478-196-3182

Query Match      23.6%; Score 51; DB 7; Length 1170;
Best Local Similarity 44.8%; Pred. No. 21;
Matches 13; Conservative 2; Mismatches 10; Indels 4; Gaps 2;

QY      10 SQCCSEDEDIFEETAAQVSPPRGKEKROW 38
DB      677 SMGASSESED--SNTASRGPP--KERKOW 701

RESULT 13
US-09-200-650E-1
; Sequence 1, Application US/09200650E
; GENERAL INFORMATION:
; APPLICANT: Patti, Joseph M.
; APPLICANT: Foster, Timothy J.
; APPLICANT: Hook, Magnus A.O.
; APPLICANT: Eldhinn, Deirdre Ni
; APPLICANT: Perkins, Samuel L.
; TITLE OF INVENTION: Extracellular Matrix-Binding Proteins from Staphylococcus aureus
; FILE REFERENCE: P06283US2/BAS
; CURRENT APPLICATION NUMBER: US/09/200,650E
; CURRENT FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,815
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: 60/098,427
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 918
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-200-650E-1

Query Match      23.1%; Score 50; DB 5; Length 918;
Best Local Similarity 23.5%; Pred. No. 22;
Matches 8; Conservative 10; Mismatches 16; Indels 0; Gaps 0;

QY      2 DDERSTDSQQCSSEDEDIFEETAAQVSPPRGKEK 35
DB      822 DSDSDSDSDSDSDSDSDSDSDSDSRVTPPPNNEQK 855

RESULT 14
US-10-273-573-8148
; Sequence 8148, Application US/10273573
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: US/10/273,573
; CURRENT FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 8148
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(45)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: July 24, 2003, 11:48:38 ; Search time 3.36164 Seconds
(without alignments)
1172.914 Million cell updates/sec

Title: US-09-991-681-28

Perfect score: 216

Sequence: 1 EDDERSTDSQQCSSEDEDI.....ETFAQVSPPRGKRWRRAR 41

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	64	29.6	707	2 A35804	nucleolin - human
2	64	29.6	712	2 JH0148	nucleolin - rat
3	63	29.2	713	2 A27441	nucleolin - Chines
4	61	28.0	2599	2 A56616	unknown protein FI
5	60.5	28.0	512	1 MMBEXA	US54 protein - hum
6	60	27.8	395	2 S66940	SGT1 protein - yea
7	59.5	27.5	206	2 S70004	hypothetical prote
8	59.5	27.5	767	2 S63182	hypothetical prote
9	59	27.3	586	2 C83262	hypothetical prote
10	59	27.3	707	1 DNMS	nucleolin - mouse
11	58.5	27.1	182	2 T46396	hypothetical prote
12	58	26.9	511	1 A48560	US54 protein - hum
13	58	26.9	512	1 MMBEXA	US54 protein - hum
14	58	26.9	796	1 A32434	abba protein - Eme
15	58	26.9	1738	2 S20614	conserved hypothec
16	58	26.9	3839	2 T19799	related to TOM1 pr
17	57.5	26.6	88	2 D95264	hypothetical prote
18	57	26.4	340	2 T51386	probable protein w
19	57	26.4	396	1 A58938	surface protein in
20	57	26.4	2643	2 T29149	hypothetical prote
21	56.5	26.2	641	2 T17278	hypothetical prote
22	56	25.9	517	2 S41081	monomane transpor
23	56	25.9	1695	2 T19823	hypothetical prote
24	56	25.9	1711	1 A47392	chromodomain-helic
25	55.5	25.7	315	2 T15460	hypothetical prote
26	55.5	25.7	351	2 T32093	hypothetical prote
27	55.5	25.7	486	2 A89927	elastin binding pr
28	55.5	25.7	1465	2 S45628	DNA-directed DNA p
29	55.5	25.7	2135	2 T14602	variant-specific s

30	55	25.5	411	2 S48647	peptidylprolyl iso
31	55	25.5	426	2 T05676	hypothetical prote
32	55	25.5	731	2 T08855	nephrocystin - hum
33	55	25.5	765	2 E96558	hypothetical prote
34	55	25.5	1040	2 S50617	BBB1 protein - yea
35	55	25.5	1241	2 JU0466	potassium transpor
36	54.5	25.2	440	2 A42136	transcription fact
37	54.5	25.2	899	2 S49634	hypothetical prote
38	54.5	25.2	1070	2 T31069	collid-BMP-1 like
39	54	25.0	206	2 S43445	translation elonga
40	54	25.0	314	2 F84886	probable C2H2-type
41	54	25.0	328	2 T01225	hypothetical prote
42	54	25.0	328	2 B84545	hypothetical prote
43	54	25.0	332	2 C40646	endospore developm
44	54	25.0	489	2 A45988	denitrin matrix acid
45	54	25.0	529	2 S44649	f42h10.7 protein -

ALIGNMENTS

RESULT 1
A35804
nucleolin - human
N:Alternate names: phosphoprotein ppi100; protein B50; protein C23
C:Species: Homo sapiens (man)
C:Date: 23-Oct-1990 #sequence revision 23-Oct-1990 #text_change 23-Jul-1999
C:Accession: A35804; S04631; A48138; A55996
R:Stivatava, M.; McBride, O.W.; Fleming, P.J.; Pollard, H.B.; Burns, A.L.
J. Biol. Chem. 265, 14922-14931, 1990
A:Title: Genomic organization and chromosomal localization of the human nucleolin gene.
A:Reference number: A35804; MUID:90368666; PMID:2394707
A:Accession: A35804
A:Molecule type: DNA
A:Residues: 1-707 <SR>
A:Cross-References: GB:06058; GB:J05584; NID:9189305; PIDN:AAA59954.1; PID:9189306
R:Stivatava, M.; Fleming, P.J.; Pollard, H.B.; Burns, A.L.
FEBS Lett. 250, 99-105, 1989
A:Title: Cloning and sequencing of the human nucleolin cDNA.
A:Reference number: S04631; MUID:89290043; PMID:2737305
A:Accession: S04631
A:Status: not compared with conceptual translation
A:Residues: 1-707 <SR2>
A:Molecule type: mRNA
R:Shikawa, F.; Matunis, M.J.; Dreyfuss, G.; Cecch, T.R.
Mol. Cell. Biol. 13, 4301-4310, 1993
A:Title: Nuclear proteins that bind the pre-mRNA 3' splice site sequence r(UUG/G) and c)
A:Reference number: A48138; MUID:93309464; PMID:8321232
A:Accession: A48138
A:Molecule type: protein
A:Residues: 458-474 <ISH>
A:Experimental source: Hela cell nuclei
A:Note: Sequence extracted from NCBI database (NCBI:134645)
R:Jordan, P.; Heid, H.; Kinzel, V.; Kuebler, D.
Biochemistry 33, 14696-14706, 1994
A:Title: Major cell surface-located protein substrates of an ecto-protein kinase are hom
A:Reference number: A55996; MUID:95086063; PMID:7993898
A:Accession: A55996
A:Molecule type: protein
A:Residues: 231-236;349-362;399-403;458-461;655-656; 'X', 658-660 <OR>
A:Experimental source: surface-labelled Hela cells
C:Genetics:
A:Gene: GDB:NCL
A:Cross-References: GDB:125908; OMIM:164035
A:Map position: 2q12-2qter
C:Superfamily: nucleolin; ribonucleoprotein repeat homology
C:Keywords: DNA binding; nucleolin; nucleus; phosphoprotein; RNA binding
F:308-373/Domain: ribonucleoprotein repeat homology <RRM1>
F:394-456/Domain: ribonucleoprotein repeat homology <RRM2>
F:487-550/Domain: ribonucleoprotein repeat homology <RRM3>
F:573-654/Domain: ribonucleoprotein repeat homology <RRM4>

Query Match 29.6%; Score 64; DB 2; Length 707;

A;Molecule type: mRNA
A;Residues: 1-713 <LAP>

C:\Date: 31-Dec-1992 #sequence_ revision 31-Dec-1992 #text_ change 16-Jun-2000
C:\Accession: JQ1498


```

Nature 406, 959-964, 2000
A>Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: C93262
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-586 <STG>
A:Cross-references: GB:AE004731; GB:AE004091; NID:g9949171; PIDN:AA06462.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA3074

Query Match      27.3% Score 59; DB 2; Length 586;
Best Local Similarity 33.3%; Pred. No. 18;
Matches 11; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

Cy 4 ERSTDSSQCCSEDEDFEETAAQVSPPRGKKR 36
Db 458 QROEAAQQQAGENKQROEASQSPSGSSOR 490

RESULT 10
DNMS
nucleolin - mouse
C:Species: Mus musculus (house mouse)
C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 22-Jun-1999
A:Accession: A29958; A40769; A56240; I84688
R:Bourbon, H.M.; Lapeyre, B.; Amalric, F.
J. Mol. Biol. 200, 627-638, 1988
A>Title: Structure of the mouse nucleolin gene. The complete sequence reveals that each
A:Reference number: A29958; MUID:88316930; PMID:3137346
A:Accession: A29958
A:Molecule type: DNA
A:Residues: 1-707 <BOU>
A:Cross-references: GB:X07699; NID:G53453; PIDN:CA30538.1; PID:G53454
R:Paternack, M.S.; Bleier, K.U.; McInerney, T.N.
J. Biol. Chem. 266, 14703-14708, 1991
A>Title: Granzyme A binding to target cell proteins. Granzyme A binds to and cleaves nucleolin
A:Reference number: A40769; MUID:91317840; PMID:1860869
A:Accession: A40769
A:Molecule type: protein
A:Residues: 2-20, 'X', 22-24 <PAS>
R:Yang, T.H.; Tsai, W.H.; Lee, Y.M.; Lei, H.Y.; Lai, M.Y.; Chen, D.S.; Yen, N.H.; Lee, S.
Mol. Cell. Biol. 14, 6066-6074, 1994
A>Title: Purification and characterization of nucleolin and its identification as a transmembrane protein
A:Reference number: A56240; MUID:94344117; PMID:8065340
A:Accession: A56240
A:Molecule type: protein
A:Residues: 2-19; 558-567 <YAN>
R:Bourbon, H.
Gene 68, 73-84, 1988
A>Title: Sequence and structure of the nucleolin promoter in rodents: Characterization of
A:Reference number: I48118; MUID:85121496; PMID:2906027
A:Accession: I84688
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-44 <RES>
A:Cross-references: GB:M20089; NID:g200111; PIDN:AAA39841.1; PID:g554246
C:Comment: This housekeeping protein is involved in the synthesis, packaging, and maturation of ribosomes
C:Genetics:
A:introns: 6/3; 45/3; 211/1; 273/1; 302/1; 349/2; 391/1; 432/3; 484/1; 525/2; 566/1; 608/1
C:Superfamily: nucleolin; ribonucleoprotein repeat homology
C:Keywords: DNA binding; duplication; nucleus; phosphoprotein; RNA binding; transcribed
F.310-375/Domain: ribonucleoprotein repeat homology <RNM1>
F.311-316/Region: RNA-binding RNP2 motif
F.344-356/Region: RNA-binding RNP2 motif
F.356-458/Domain: ribonucleoprotein repeat homology <RRM2>
F.397-402/Region: RNA-binding RNP2 motif
F.431-438/Region: RNA-binding RNP1 motif
F.488-551/Domain: ribonucleoprotein repeat homology <RRM3>
F.488-494/Region: RNA-binding RNP2 motif
F.552-531/Region: RNA-binding RNP1 motif
F.570-634/Domain: ribonucleoprotein repeat homology <RRM4>

```

```

F:571-576/Region: RNA-binding RNP2 motif
F:607-614/Region: RNA-binding RNP1 motif

Query Match          27.3%;   Score 59;   DB 1;   Length 707;
Best Local Similarity 32.4%;   Pred. No. 22;
Matches 11; Conservative 7; Mismatches 16; Indels 0; Gaps 0;

Oy      1 EDDERSTDSQCCSDEDFEETAOVSPRGKE 34
        ||| | | | | | | | | | | | | | | |
Db      193 EDDEDDEEDDEDEDEDEDESEEVMEITTKAGKK 226

RESULT 11
T46396
hypothetical protein DKFPz43411820.1 - human (Fragment)
C:Species: Homo sapiens (man)
C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C:Accession: T46396
R:Ottenwaelder, B.; Obermaier, B.; Wewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A:Reference number: Z23031
A:Accession: T46396
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-182 <AA>
A:Cross-references: EMBL:AL137730
A:Experimental source: adult testis; clone DKFPz43411820
C:Genetics:
A>Note: DKFPz43411820.1

Query Match          27.1%;   Score 58.5;   DB 2;   Length 182;
Best Local Similarity 34.8%;   Pred. No. 6.1;
Matches 16; Conservative 8; Mismatches 17; Indels 5; Gaps 1;

Oy      1 EDDERSTDSQCCSDEDE-----IFEETAOVSPPRGKRRQWRAR 41
        ||| | | | | | | | | | | | | | | |
Db      11 EDELISSESESYESTDDEDRGRMVKLMELANLPKPKRTIKQHVR 56

RESULT 12
A48560
UL54 protein - human herpesvirus 1 (strain HFEM)
C:Species: human herpesvirus 1
C>Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 16-Feb-1997
C:Accession: A48560
R:Rozen-Wolfe, A.; Frank, S.; Raab, K.; Moyal, M.; Becker, Y.; Darai, G.
Virus Res. 25, 189-199, 1992
A>Title: Determination of the coding capacity of the BamHI DNA fragment B of apathogenic
A:Reference number: A48560; MUID:93070559; PMID:1332274
A:Accession: A48560
A:Molecule type: DNA
A:Residues: 1-511 <ROS>
A>Note: sequence extracted from NCBI backbone (NCBIN:117573, NCBIP:117574)
C:Genetics:
A:Gene: UL54
C:Superfamily: varicella-zoster virus gene 4 protein
C:Keywords: transcription regulation

Query Match          26.9%;   Score 58;   DB 1;   Length 511;
Best Local Similarity 40.7%;   Pred. No. 21;
Matches 11; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

Oy      3 DERSTDSQCCSDEDFEETAOVSP 29
        ||| | | | | | | | | | | | | | | |
Db      33 DDLESDSGECSSSDMEDPHGEDGP 59

RESULT 13
WMBEX4
UL54 protein - human herpesvirus 1 (strain 17)
C:Species: human herpesvirus 1
C>Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 16-Jun-2000
C:Accession: I30089
```



```

R:McGeoch, D.U.; Dalmaypale, M.A.; Davison, A.J.; Dolan, A.; Frame, M.C.; McNab, D.; Per
J. Gen. Virol. 69, 1511-1574, 1988
A:Title: The complete DNA sequence of the long unique region in the genome of herpes sim
A:Reference number: A30083; MUID:88274327; PMID:2835594
A:Accession: 130083
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-512 <MCG>
A:Cross-references: GB:X14112; NID:g1944536; PIDN:CAA32290.1; PID:g59554; GB:D00317
C:Genetics:
A:Gene: UL54
C:Superfamily: varicella-zoster virus gene 4 protein
C:Keywords: transcription regulation

Query Match          26.9%  Score 58;  DB 1;  Length 512;
Best Local Similarity 40.7%  Pred. No. 21;
Matches 11; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

Cy      3  DERSTDSQCCSDEDDIFFEETAQSP 29
      ||:|||||:|||||:|
Db      33  DDLSDSSGCCSSDEDDPHGEDGP 59

```

```

A:Molecule type: DNA
A:Residues: 1-1738 <WOM>
A:Cross-references: EMBL:M81884, NID:G336917, PID:G336940
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1992
C:Genetics:
A:Genome: plastid
C:Keywords: plastid

Query Match          26.9%  Score 58 ; DB 2;  length 1738;
Best local similarity 48.4%  Pred. No. 75;
Matches 15;  conservative 2;  Mismatches 12;  Indels 2;  Gaps 1;

QY      7  TDSQCCSEDEDIETETAQVSPRCKEKQ 37
          :| ||||| || |||||
Db      245  TLKKEQKQVEDEDI--ETVNMQMPGLGKRE 273

Search completed: July 24, 2003, 11:58:02
Job time : 7.35164 secs

```

```

RESULT 14
A32434
abaa protein - Emerizella nidulans
C:Species: Emerizella nidulans, Aspergillus nidulans
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A32434
R:Mirabito, P.M.; Adams, T.H.; Timberlake, W.E.
Cell 57, 859-868, 1989
A:Title: Interactions of three sequentially expressed genes control temporal and spatial
A:Reference number: A32434; MUID:89249350; PMID:2655931
A:Accession: A32434
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-796 <MIR>
A:Cross-references: GB:J04850; NID:g167997; PIDN:AAA33286.1; PID:g167998
C>Note: The authors translated the codon ATC for residue 641 as Thr
C:Superfamily: abaa protein; TEA DNA-binding domain homology
C:Keywords: DNA binding; nucleus; transcription regulation
F,131-201/Domain: TEA DNA-binding domain homology <TEA>

Query Match      26.9%   Score 58;   DB 1;   Length 796;
Best Local Similarity 35.3%;   Pred. No. 33;
Matches 12; Conservative 6; Mismatches 14; Indels 2; Gaps 1;

Oy      8  DSSQCSSEDEDPFETAOVSPPRGEXKRWRR 41
      | | | | | : : : | : | : |
Db      135 DGPVWSDELEDAFOQALEANPPMG--RRKWSR 166

```

RESULT 15
S20614
conserved hypothetical protein 1738 - beechdrops plastid
C:Species: plastid Epilafus virginiana (beechdrops)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 05-Jun-1998
C:Accession: S20614; S78401; S16720
R:Wolfe, K.H.; Morden, C.W.; Palmer, J.D.
J. Mol. Biol. 223, 95-104, 1992
A:Title: Small single-copy region of plastid DNA in the non-photosynthetic angiosperm Ep
on at a non-bioenergetic locus
A:Reference number: S20614; MUID:92114084; PMID:1731088
A:Accession: S20614
A:Molecule type: DNA
A:Residues: 1-1738 <MOL>
A:Cross-references: EMBL:X61368; NID:g11550; PID:g11551
R:Wolfe, K.H.; Morden, C.W.; Ems, S.C.; Palmer, J.D.
J. Mol. Evol. 35, 304-317, 1992
A:Title: Rapid evolution of the plastid translational apparatus in a nonphotosynthetic E
A:Reference number: S78378; MUID:93021155; PMID:1404416
A:Accession: S78401
A:Status: nucleic acid sequence not shown; translation not shown

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 11:46:57 ; Search time 1.86091 Seconds
(without alignments)
1036.105 Million cell updates/sec

Title: US-09-991-681-28

Perfect score: 216
Sequence: 1 EDDERSTDSQCCSEDEDI.....ETTAQVSPRGRKRWRRAR 41

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 10%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	64	29.6	706	1 NUCL_HUMAN	P13338 homo sapien
2	64	29.6	712	1 NUCL_RAT	P13383 ratius norv
3	63	29.2	713	1 NUCL_MESAU	P08199 mesocricetu
4	62	28.7	2799	1 EDD_HUMAN	O09071 homo sapien
5	60.5	28.0	512	1 IE63_HSVZH	P28276 herpes simp
6	60	27.8	395	1 SGT1_YEAST	O08446 saccharomyc
7	59.5	27.5	767	1 YNM4_YEAST	P35866 saccharomyc
8	59	27.3	706	1 NUCL_MOUSE	P03405 mus musculu
9	58.5	27.1	566	1 CC45_HUMAN	O75419 homo sapien
10	58.5	27.1	582	1 NMT_HUMAN	O09583 homo sapien
11	58.5	27.1	582	1 PEGC_BRARE	P79741 brachydantio
12	58	26.9	511	1 IE63_HSVIE	P16295 herpes simp
13	58	26.9	512	1 IE63_HSV11	P10238 herpes simp
14	58	26.9	796	1 ABBA_EBENI	P20945 emericella
15	58	26.9	1738	1 YCF1_EPIVI	O00383 epifagus vi
16	57	26.4	597	1 VBI_COPV	O08536 canine oral
17	56.5	26.2	580	1 RGP1_XENLA	O13066 xenopus lae
18	56.5	26.2	692	1 ANR6_HUMAN	O09294 homo sapien
19	56	25.9	517	1 VMT2_BOVIN	O21963 bos taurus
20	56	25.9	522	1 PLIN_HUMAN	O60240 homo sapien
21	56	25.9	1711	1 CHD1_MOUSE	P40201 mus musculu
22	55.5	25.7	411	1 MP62_LYTP1	P91753 lyechinus
23	55.5	25.7	618	1 ORC2_DROME	O24168 drosophila
24	55.5	25.7	851	1 CT06_HUMAN	O98501 homo sapien
25	55.5	25.7	1465	1 DPOA_MOUSE	P33609 mus musculu
26	55	25.5	147	1 DPB3_HUMAN	O98719 homo sapien
27	55	25.5	411	1 FRB3_YEAST	P36911 saccharomyc
28	55	25.5	503	1 SSG1_MOUSE	O31918 mus musculu
29	55	25.5	732	1 DMPI_MOUSE	O51588 mus musculu
30	55	25.5	910	1 NPH1_HUMAN	O15259 homo sapien
31	55	25.5	910	1 US1C_MOUSE	O96864 mus musculu
32	55	25.5	1040	1 BO12_YEAST	P33969 saccharomyc
33	55	25.5	1241	1 TRK1_SACBA	P28569 saccharomyc

34	54.5	25.2	440	1 SCAL_DROME	P30052 drosophila
35	54.5	25.2	899	1 YMJ3_YEAST	O04500 saccharomyc
36	54.5	25.2	1282	1 BMS1_HUMAN	O14692 homo sapien
37	54	25.0	205	1 EP1B_YEAST	P32471 saccharomyc
38	54	25.0	332	1 SP2B_BACSU	P37575 bacillus bu
39	54	25.0	489	1 DMPI_RAT	P98159 ratius norv
40	54	25.0	529	1 DGR6_CAEEL	P34420 caenorhabdi
41	54	25.0	677	1 UBFI_XENLA	P35979 xenopus lae
42	54	25.0	699	1 SRCH_HUMAN	P23327 homo sapien
43	54	25.0	1085	1 IFH1_YEAST	P39520 saccharomyc
44	54	25.0	1912	1 VITI_CHICK	P87498 gallus gall
45	54	25.0	2414	1 P300_HUMAN	O09472 homo sapien

ALIGNMENTS

```

RESULT 1
NUCL_HUMAN STANDARD; PRT; 706 AA.
ID NUCL_HUMAN
AC P13338;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Nucleolin (Protein C23).
GN NCL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=69290043; PubMed=2737305;
RA Srivastava M., Fleming P.J., Pollard H.B., Burns A.L.;
RT "Cloning and sequencing of the human nucleolin cDNA.";
RL FEBS Lett. 250:99-105(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90368666; PubMed=2394707;
RA Srivastava M., McBride O.W., Fleming P.J., Pollard H.B., Burns A.L.;
RT "Genomic organization and chromosomal localization of the human
nucleolin gene.";
RL J. Biol. Chem. 265:14922-14931(1990).
CC -!- FUNCTION: NUCLEOLIN IS THE MAJOR NUCLEOLAR PROTEIN OF GROWING
EUKARYOTIC CELLS. IT IS FOUND ASSOCIATED WITH INTRANUCLEOLAR
CHROMATIN AND PRERIBOSOMAL PARTICLES. IT INDICES CHROMATIN
DECONDENSATION BY BINDING TO HISTONE H1. IT IS THOUGHT TO PLAY A
ROLE IN PRE-RRNA TRANSCRIPTION AND RIBOSOME ASSEMBLY.
CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar.
CC -!- SIMILARITY: Contains 4 RNA recognition motif (RRM) domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
or send an email to license@ebi.ac.uk).
CC -----
CC EMBL; M60858; AAAS9954.1; -.
CC PIR; A35804; A35804.
CC HSSP; P09651; 1HA1.
CC Aarhus/Ghent-2DPAGE; 1210; NEPHGE.
CC Genew; HGNC:7667; NCL.
CC MIM; 164035; -.
CC GO; GO:0005730; C:nucleolus; TAS.
CC GO; GO:0003723; F:RNA binding activity; TAS.
CC InterPro; IPR000504; RNA_rec_mot.
CC Pfam; PF00076; rrm; 4.
CC SMART; SM00360; RRM; 4.
CC PROSITE; PS50102; RRM; 4.
CC PROSITE; PS00030; RRM_RNP_1; 3.
CC Nuclear protein; Phosphorylation; Methylation; DNA-binding; Repeat;
KW

```

```

KW RNA-binding.
FT INIT MET 0
FT DOMAIN 142 170
FT DOMAIN 184 208
FT DOMAIN 233 270
FT DOMAIN 306 382
FT DOMAIN 392 465
FT DOMAIN 485 559
FT DOMAIN 571 643
FT DOMAIN 645 694
FT DOMAIN 57 134
FT REPEAT 57 64
FT REPEAT 74 81
FT REPEAT 82 89
FT REPEAT 90 97
FT REPEAT 98 103
FT REPEAT 104 111
FT REPEAT 119 126
FT REPEAT 127 134
FT MOD_RES 144 152
FT MOD_RES 152 183
FT MOD_RES 183 183
SQ SEQUENCE 706 AA; 76213 MW; 85A2FC2CA22EA03DB CRC64;

Query Match
Best Local Similarity 29.6%; Score 64; DB 1; Length 706;
Matches 12; Conservative 6; Mismatches 16; Indels 0; Gaps 0;

1 EDDERSTDSQCCSEDEDIFETTAQVSPPRGKE 34
186 EDDDEDEDEDEDDDEDEDESEEMETTPAKGK 219

RESULT 2
NUCL_RAT STANDARD; PRT; 712 AA.
AC P13383;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Nucleolin (Protein C23).
GN NCL OR NUC.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90269607; PubMed=2347493;
RA "Bourdon H.-M., Amalric F.;
RT "Nucleolin gene organization in rodents: highly conserved sequences
RL within three of the 13 introns.";
RL Gene 88:187-196(1990).
RN [2]
RP SEQUENCE OF 1-44 FROM N.A.
RX MEDLINE=89121496; PubMed=2906027;
RA Bourdon H.-M., Prudhomme M., Amalric F.;
RT "Sequence and structure of the nucleolin promoter in rodents:
RL characterization of a strikingly conserved Cpg island.";
RL Gene 68:73-84(1988).
RN [3]
RP FUNCTION: NUCLEOLIN IS THE MAJOR NUCLEOLAR PROTEIN OF GROWING
CC EUKARYOTIC CELLS. IT IS FOUND ASSOCIATED WITH INTRANUCLEOLAR
CC CHROMATIN AND PERIBIOSOMAL PARTICLES. IT INDUCES CHROMATIN
CC RECONDENSATION BY BINDING TO HISTONE H1. IT IS THOUGHT TO PLAY A
CC ROLE IN PRE-RNA TRANSCRIPTION AND RIBOSOME ASSEMBLY.
CC -1- SUBCELLULAR LOCATION: Nuclear; nucleolar.
CC -1- SIMILARITY: Contains 4 RNA recognition motif (RRM) domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isd-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL, M55022; AAA41732.1; -.
DR EMBL, M55015; AAA41732.1; JOINED.
DR EMBL, M55017; AAA41732.1; JOINED.
DR EMBL, M55020; AAA41732.1; JOINED.
DR EMBL, M20900; AAA41733.1; -.
DR HSSP, P09651; 1HA1.
DR InterPro, IPR000504; RNA_rec_mot.
DR Pfam, PF00076; rrm; 4.
DR SMART, SM00360; RRM; 4.
DR PROSITE, PS50102; RRM; 4.
DR PROSITE, PS00030; RRM_RNP_1; 3.
KW Nuclear protein; Phosphorylation; Methylation; DNA-binding; Repeat;
KW RNA-binding.
FT INIT MET 0
FT DOMAIN 142 167
FT DOMAIN 187 215
FT DOMAIN 241 274
FT DOMAIN 310 386
FT DOMAIN 396 469
FT DOMAIN 488 562
FT DOMAIN 574 649
FT DOMAIN 651 702
FT DOMAIN 57 134
FT REPEAT 57 64
FT REPEAT 74 81
FT REPEAT 82 89
FT REPEAT 90 97
FT REPEAT 98 103
FT REPEAT 104 111
FT REPEAT 119 126
FT REPEAT 127 134
SQ SEQUENCE 712 AA; 77016 MW; 68774A214E550F90 CRC64;

Query Match
Best Local Similarity 29.6%; Score 64; DB 1; Length 712;
Matches 10; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

1 EDDERSTDSQCCSEDEDIFETTAQVSPPRGKE 34
193 DDDDEDDDDDEDEDEDESEEMETTPAKGK 226

RESULT 3
NUCL_MESAU STANDARD; PRT; 713 AA.
AC P08199;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Nucleolin (Protein C23).
GN NCL.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxId=10036;
RN [1]
RP SEQUENCE OF 35-713 FROM N.A., AND SEQUENCE OF 1-34.
RX MEDLINE=87175501; PubMed=3470736;
RA Lapeyre B., Bourdon H., Amalric F.;
RT "Nucleolin, the major nucleolar protein of growing eukaryotic cells:
RL an unusual protein structure revealed by the nucleotide sequence.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:1472-1476(1987).
RN [2]
RP FUNCTION.
RX MEDLINE=88312631; PubMed=3409881;
RA Errard M.S., Belenguer P., Calzergues-Ferrer M., Pantaloni A.,
RA Amalric F.;

```

"A major nucleolar protein, nucleolin, induces chromatin decondensation by binding to histone H1.";
 RL Eur. J. Biochem. 175:525-530(1988).
 CC -1- FUNCTION: NUCLEOLIN IS THE MAJOR NUCLEOLAR PROTEIN OF GROWING
 CC EUKARYOTIC CELLS. IT IS FOUND ASSOCIATED WITH INTRANUCLEOLAR
 CC CHROMATIN AND PERIBIOSOMAL PARTICLES. IT INDUCES CHROMATIN
 CC DECONDENSATION BY BINDING TO HISTONE H1. IT IS THOUGHT TO PLAY A
 CC ROLE IN PRE-RNA TRANSCRIPTION AND RIBOSOME ASSEMBLY.
 CC -1- SUBCELLULAR LOCATION: Nuclear; nucleolar.
 CC -1- SIMILARITY: Contains 4 RNA recognition motif (RRM) domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M15825; AAA36966.1; -;
 DR PIR; A27441; A27441.
 DR PDB; 1EJ7; 16-OCT-00.
 DR PDB; 1EJC; 16-OCT-00.
 DR PDB; 1EJE; 03-JAN-01.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF00076; rrm; 4.
 DR SMART; SMO0360; RRM; 4.
 DR PROSITE; PS00030; RRM_RNP_1; 3.
 DR Nucleic acid binding; Phosphorylation; Methylation; DNA-binding; Repeat;
 KW RNA-binding; 3D-structure.
 FT INTR MET 0 0
 FT DOMAIN 141 169 ASP/GLU-RICH (ACIDIC).
 FT 188 213 ASP/GLU-RICH (ACIDIC).
 FT DOMAIN 239 271 ASP/GLU-RICH (ACIDIC).
 FT 307 363 RNA-BINDING (RRM) 1.
 FT DOMAIN 393 466 RNA-BINDING (RRM) 2.
 FT 485 559 RNA-BINDING (RRM) 3.
 FT DOMAIN 571 646 RNA-BINDING (RRM) 4.
 FT 648 701 ARG/GLY/PHE-RICH.
 FT DOMAIN 133 133 8 X 8 AA TANDEM REPEATS OF X-T-P-X-K-K-X-
 FT X-X.
 FT REPEAT 56 63 1.
 FT REPEAT 73 80 2.
 FT REPEAT 81 88 3.
 FT REPEAT 89 96 4.
 FT REPEAT 97 102 5 (INCOMPLETE).
 FT REPEAT 103 110 6.
 FT REPEAT 118 125 7.
 FT REPEAT 126 133 8.
 FT REPEAT 217 221 X-T-P-X-K-K-X-X MOTIF.
 FT MOD_RES 143 143 PHOSPHORYLATION.
 FT MOD_RES 156 156 PHOSPHORYLATION.
 FT MOD_RES 187 187 PHOSPHORYLATION.
 FT MOD_RES 655 655 METHYLATION (DI-).
 FT MOD_RES 659 659 METHYLATION (DI-).
 FT MOD_RES 665 665 METHYLATION (DI-).
 FT MOD_RES 669 669 METHYLATION (DI-).
 FT MOD_RES 673 673 METHYLATION (DI-).
 FT MOD_RES 679 679 METHYLATION (DI-).
 FT MOD_RES 681 681 METHYLATION (DI-).
 FT MOD_RES 687 687 METHYLATION (DI-).
 FT MOD_RES 691 691 METHYLATION (DI-).
 FT MOD_RES 694 694 METHYLATION (DI-).
 SQ SEQUENCE 713 AA; 76997 MW; 79DDCF724CED7DB4 CRC64;
 Query Match 29.2%; Score 63; DB 1; Length 713;
 Best Local Similarity 32.4%; Pred. No. 4.5;
 Matches 12; Conservative 8; Mismatches 17; Indels 0; Gaps 0;

RESULT 4
 EDD HUMAN STANDARD; PRT; 2799 AA.
 AC 095071; 094970; Q9NPL3;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Ubiquitin-protein ligase EDD (EC 6.3.2.-) (Hyperplastic discs
 DE protein homolog) (hHYD) (Progesterin-induced protein).
 GN EDD OR HYD OR KIA0896.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eumetazoa; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Heart, and Placenta;
 RX MEDLINE=99153743; PubMed=10030672;
 RA Callaghan M.J., Russel A.J., Woollett E., Sutherland G.R.,
 RA Sutherland R.L., Watts C.K.W.;
 RT "Identification of a human HECT family protein with homology to the
 RT Drosophila tumor suppressor gene hyperplastic discs.";
 RL Oncogene 17:3479-3491(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal brain;
 RX MEDLINE=21671350; PubMed=11714696;
 RA Honda Y., Tojo M., Matsuzaki K., Anan T., Matsumoto M., Ando M.,
 RA Saito H., Nakao M.;
 RT "Cooperation of HECT-domain ubiquitin ligase hHYD and DNA
 RT topoisomerase II-binding protein for DNA damage response.";
 RL J. Biol. Chem. 277:3599-3605(2002).
 RN [3]
 RP SEQUENCE OF 714-2799 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22158633; PubMed=12168954;
 RA Nakajima D., Okazaki N., Yamakawa H., Kikuno R., Ohara O., Nagase T.;
 RT "Construction of expression-ready cDNA clones for KIAA genes: manual
 RT curation of 330 KIAA cDNA clones.";
 RL DNA Res. 9:99-106(2002).
 RN [4]
 RP SEQUENCE OF 1569-2799 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=99156230; PubMed=10048485;
 RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirosewa M.,
 RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XII.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro.";
 RL DNA Res. 5:355-364(1998).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (1.04 ANGSTROMS) OF 2391-2455, AND MUTAGENESIS
 RP OF CYS-2768.
 RX MEDLINE=21192643; PubMed=11287654;
 RA Deo R.C., Sorenberg N., Buxley S.K.;
 RT "X-ray structure of the human hyperplastic discs protein: an ortholog
 RT of the C-terminal domain of poly(A)-binding protein.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4414-4419(2001).
 CC -1- FUNCTION: E3 ubiquitin-protein ligase which accepts ubiquitin from
 CC an E2 ubiquitin-conjugating enzyme in the form of a thioester and
 CC then directly transfers the ubiquitin to targeted substrates (By
 CC similarity). This protein may be involved in maturation and/or
 CC post-transcriptional regulation of mRNA. May play a role in
 CC control of cell cycle progression. May have tumor suppressor
 CC function. Regulates DNA topoisomerase II binding protein (TopBP1)
 CC for the DNA damage response.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- TISSUE SPECIFICITY: Widely expressed. Most abundant in testis and
 CC expressed at high levels in brain, pituitary and kidney.
 CC -1- MISCELLANEOUS: A cysteine residue is required for ubiquitin-
 CC thioester formation.

```

CC -! SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
CC domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF006010; AAD01259.2; -
CC EMBL: U95000; AAF88143.1; -
CC EMBL: AB020703; BAA74919.2; -
CC PDB: 1I2T; 18-APR-02.
CC GO: GO:0005625; C:soluble fraction; TAS.
CC GO: GO:0004840; F:ubiquitin conjugating enzyme activity; TAS.
CC GO: GO:0004842; F:ubiquitin-protein ligase activity; TAS.
CC GO: GO:0008283; P:cell proliferation; TAS.
CC InterPro: IPR000569; HECT domain.
CC InterPro: IPR002004; PABP/HECT.
CC InterPro: IPR000449; UBA domain.
CC InterPro: IPR003126; Znf_Nrecoznln.
CC Pfam: PF00658; HECT; 1.
CC Pfam: PF00627; UBA; 1.
CC Pfam: PF02007; zf-UBR1; 1.
CC SMART: SM00113; HECTC; 1.
CC SMART: SM00517; POLYA; 1.
CC SMART: SM00396; Znf_UBR1; 1.
CC PROSITE: PS50237; HECT; 1.
CC Ubl conjugation pathway; Ligase: Nuclear protein; 3D-structure.
CC -----
FT DOMAIN 2353 2449 PABP-LIKE.
FT 2462 2799 HECT.
FT 1986 1997 ASP/GLU-RICH (ACIDIC).
FT 2036 2059 PRO-RICH.
FT 2329 2348 ARG/GLU-RICH (MIXED CHARGE).
FT 2357 2366 ARG/ASP-RICH (MIXED CHARGE).
FT 2489 2500 ASP/GLU-RICH (ACIDIC).
FT 2737 2757 PRO-RICH.
FT 940 945 POLY-GLU.
FT 985 985 POLY-SER.
FT 1528 1537 POLY-SER.
FT 1671 1681 POLY-SER.
FT 1762 1768 POLY-ALA.
FT 2768 2768 UBIQUITIN (BY SIMILARITY).
FT 2768 2768 C->A: LOSS OF UBIQUITIN BINDING.
FT 134 134 S-> P (IN REF. 2).
FT 229 229 E-> K (IN REF. 2).
FT 258 258 S-> Y (IN REF. 2).
FT 374 375 IG-> M (IN REF. 2).
FT 772 772 D-> H (IN REF. 2).
FT 780 780 Q-> R (IN REF. 2).
FT 884 884 D-> G (IN REF. 2).
FT 1811 1811 S-> P (IN REF. 2).
FT 2144 2144 L-> H (IN REF. 2).
FT 2282 2282 K-> R (IN REF. 2).
FT 2474 2474 MISSING (IN REF. 3 AND 4).
FT 2489 2489 D-> N (IN REF. 2).
SQ SEQUENCE 2799 AA; 309348 MW; 871300DB404FF561 CRC64;

Query Match 28.7%; Score 62; DB 1; Length 2799;
Best Local Similarity 29.8%; Pred. No. 27;
Matches 17; Conservative 6; Mismatches 18; Indels 16; Gaps 2;

```

```

AC P28276;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Transcriptional regulator IE63 (VIMW63) (ICP27).
GN UL54.
OS Herpes simplex virus (type 2 / strain HG52).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10315;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92113549; PubMed=1662697;
RA McGeoch D.J., Cunningham C., McIntyre G., Dolan A.;
RT "Comparative sequence analysis of the long repeat regions and
RT adjoining parts of the long unique regions in the genomes of herpes
RT simplex viruses types 1 and 2."
RL J. Gen. Virol. 72:3057-3075 (1991).
RN [2]
RP SEQUENCE FROM N.A.
RA Dolan A.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
CC -! FUNCTION: INVOLVED IN THE DOWN-REGULATION OF VIRAL IMMEDIATE-EARLY
CC GENES AND ACTS IN COMBINATION WITH ICP0 AND ICP4 AS AN ACTIVATOR
CC OF LATE GENES (BY SIMILARITY).
CC -! SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL54,
CC HSV-2 UL54, EBV-1 5, VZV 4, EBV BMLF1, HCMV UL69, AND HVS-1 57.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D10471; BAA01269.1; -
DR EMBL: Z86099; CAB06702.1; -
DR PIR: J01498; WMBEXA.
KW Early protein; Transcription regulation; Activator; DNA-binding.
SQ SEQUENCE 512 AA; 54958 MW; 459651470A503BA7 CRC64;

Query Match 28.0%; Score 60.5; DB 1; Length 512;
Best Local Similarity 34.1%; Pred. No. 6.2;
Matches 14; Conservative 7; Mismatches 13; Indels 7; Gaps 1;

```

```

QY 1 EDDERSTDS--SQCCSSEDEDFEETAVSP-----RGEKERQMPAR 41
ID 1E63_HSV2H 1663 EDDSGSNSSDSSSSQSDDIETFMDEPLETNTSSHANGAQAQPRSMQMAVR 1719
ID 1E63_HSV2H STANDARD; PRT; 512 AA.

```

```

RESULT 6
ID SGT1_YEAST STANDARD; PRT; 395 AA.
AC 008446;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE SGT1 Protein.
GN SGT1 OR YOR057W OR YOR29-08.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Kitagawa K., Connolly C., Hieter P.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97279235; PubMed=9133743;
RA Valens M., Bohn C., Daignan-Fornier B., Dang V., Bolotin-Fukuhara M.;
RT "The sequence of a 54.7 kb fragment of yeast chromosome XV reveals

```



```

RT "Characterization of CDC45L: a gene in the 22q11.2 deletion region
RT expressed during murine and human development.";
RL Mamm. Genome 10:322-326 (1999).
RP
RX SEQUENCE FROM N.A.
RA MEDLINE=98391766; PubMed=9724329;
RA McKie J.M., Madley R.B., Sutherland H.F., Taylor C.L., Scambler P.J.;
RT "Direct selection of conserved CNAs from the DiGeorge critical
RT region: isolation of a novel CDC45-like gene.";
RL Genome Res. 8:834-841 (1998).
CC -1- FUNCTION: REQUIRED FOR INITIATION OF CHROMOSOMAL DNA REPLICATION.
CC -1- SUBUNIT: ASSOCIATED WITH ORC2L.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear.
CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED, HIGHEST LEVELS ARE FOUND IN
CC ADULT TESTIS AND TYMUS AND IN FETAL LIVER.
CC -1- DEVELOPMENTAL STAGE: TRANSCRIPT PEAKS AT G1-S TRANSITION, BUT
CC TOTAL PROTEIN REMAINS CONSTANT THROUGHOUT THE CELL CYCLE.
CC EXPRESSED IN MULTIPLE TISSUES DURING EMBRYOGENESIS, INCLUDING
CC NURAL CREST-DERIVED STRUCTURES.
CC -1- SIMILARITY: BELONGS TO THE CDC45 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/submit/
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL, AF053074, AAC27289.1, -.
DR EMBL, AF081535, AAC08998.1, -.
DR EMBL, AJ223728, CAA11530.1, -.
DR Geneml, HGNC:1739, CDC45L.
DR GK, O75419, -.
DR MIM, 603465, -.
DR GO, GO:0005634, C:nucleus; TAS.
DR GO, GO:0000076, P:DNA replication checkpoint; TAS.
DR GO, GO:0006270, P:DNA replication initiation; TAS.
DR InterPro, IPR003874, CDC45_like.
DR Pfam, PF02724, CDC45, 1.
DR DNA replication; Cell cycle; Nuclear protein.
FT CONFLICT 115 115 I -> V (IN REF. 3).
SQ SEQUENCE 566 AA; 65568 MW; AE1BE2C8C85FE867 CRC64;

Query Match 27.1%; Score 58.5; DB 1; Length 566;
Best Local Similarity 29.8%; Pred. No. 12;
Matches 14; Conservative 13; Mismatches 9; Indels 11; Gaps 2;

QY 1 EDDERS-----TDSGQSGSEDDIDFEETRAQVSPRGKREKRRAR 41
   ||| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 139 EDEHSGNDSGDGSEPSSEKRTREEIVEQTM-----RRRQREWEAR 180

RESULT 10
MNT_HUMAN
ID MNT_HUMAN STANDARD; PRT; 562 AA.
AC O99583;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Max binding protein MNT (ROX protein) (MYC antagonist MNT).
DE MNT OR ROX.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=97327566; PubMed=9184233;
RA Meroni G., Raymond A., Alcalay M., Borani G., Tanigami A.,
RA Tonlorenzi R., Lo Nigro C., Messali S., Zollo M., Ledbetter D.H.,
RA Brent R., Ballabio A., Carrozzo R.;
```


RT "Box, a novel bHLHZip protein expressed in quiescent cells that
RT heterodimerizes with Max, binds a non-canonical E box and acts as a
RT transcriptional repressor."
RL EMBL J. 16:2892-2906(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98260677; PubMed=9598315;
RA Nigro C.L., Venesio T., Raymond A., Meroni G., Alberici P.,
RA Caimarca S., Enrico F., Stack M., Ledbetter D.H., Lisaia D.S.,
RA Ballabio A., Carozzo R.,
RT "The human ROX Gene: genomic structure and mutation analysis in human
RT breast tumors."
RL Genomics 49:275-282(1998).
CC -1- FUNCTION: BINDS DNA AS A HETERODIMER WITH MAX AND REPRESSES
CC TRANSCRIPTION. BINDS TO THE CANONICAL E BOX SEQUENCE
CC 5'-CACGCG-3' AND, WITH HIGHER AFFINITY, TO 5'-CACGCG-3'.
CC -1- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER
CC BHLH PROTEIN. BINDS DNA AS AN HOMODIMER OR A HETERODIMER WITH MAX.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS. BHLH-ZIP SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X96401; CAA65265.1; -.
DR EMBL; Y13440; CAA73851.1; -.
DR EMBL; Y13441; CAA73851.1; JOINED.
DR EMBL; Y13442; CAA73851.1; JOINED.
DR EMBL; Y13443; CAA73851.1; JOINED.
DR EMBL; Y13444; CAA73851.1; JOINED.
DR HSSP; P25912; 1HLH.
DR TRANSFAC; T03268; -.
DR Genew; HGNC:7188; MNT.
DR MIM; 603039; -.
DR GO; GO:0003713; P:transcription co-activator activity; TAS.
DR GO; GO:0003714; P:transcription co-repressor activity; TAS.
DR GO; GO:0003700; P:transcription factor activity; TAS.
DR GO; GO:0007275; P:development; TAS.
DR GO; GO:0008285; P:negative regulation of cell proliferation; TAS.
DR GO; GO:0008074; P:regulation of cell cycle; TAS.
DR GO; GO:0006386; P:transcription from Pol II promoter; TAS.
DR InterPro; IPR01092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS00038; HLH_2; 1. FALSE_NEG.
DR PROSITE; PS50888; HLH_2; 1.
KM Transcription regulation; Repressor; Nuclear protein; DNA-binding.
FT DNA_BIND 222 233 BASIC DOMAIN.
FT DOMAIN 234 270 HELIX-LOOP-HELIX MOTIF (POTENTIAL).
FT DOMAIN 271 299 LEUCINE-ZIPPER.
SQ SEQUENCE 582 AA; 62239 MW; 06AC320D79BF18A0 CRC64;

Query Match 27.1%; Score 58.5; DB 1; Length 582;
Best Local Similarity 34.2%; Pred. No. 12;
Matches 13; Conservative 9; Mismatches 15; Indels 1; Gaps 1;

OY 1 EDDERSTDSQOCSS-DEDFEETAQVSPRGKEKQ 37
Db 323 EDDQASTSTASGEDNIDEDMEEDRAGIGLGPXKSHRPQ 360

RESULT 11
PESC_BRARE STANDARD; PRT; 582 AA.
AC P79741; Q91906;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Pescadillo.
GN PES.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97138157; PubMed=8985183;
RA Allende M.L., Amsterdam A., Becker T., Kawakami K., Galano N.,
RA Hopkins N.,
RT "Insertional mutagenesis in zebrafish identifies two novel genes,
RT pescadillo and dead eye, essential for embryonic development.";
RL Genes Dev. 10:3141-3155(1996).
RN [2]
RP SEQUENCE OF 1-180 FROM N.A.
RA Kawakami K., Grosheans H., Hopkins N.,
RT "The genomic sequence of the zebrafish pescadillo gene."
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: May have an essential role for embryonic development.
CC -1- SUBCELLULAR LOCATION: Nuclear; nucleolar.
CC -1- DEVELOPMENTAL STAGE: Widely and highly expressed during the
CC first 3 days of embryogenesis. Prominent sites of expression are
CC the eyes and optic tectum on day 1, the fin buds, liver
CC primordium, and gut on day 2, and the branchial arches on day 3.
CC -1- MISCELLANEOUS: PES mutant embryos exhibit smaller eyes, a reduced
CC brain and visceral skeleton, shortened fins and a lack of
CC expansion of liver and gut, and die on day 6 of development.
CC -1- SIMILARITY: Contains 1 BRCT domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U77627; AAB61138.1; -.
DR EMBL; AB046115; BAB01764.1; -.
DR EMBL; AB046108; BAB01764.1; JOINED.
DR ZFIN; ZDB-GENE-990415-206; pes.
DR InterPro; IPR001357; BRCT.
DR Pfam; PF00533; BRCT; 1.
DR SMART; SM00292; BRCT; 1.
DR PROSITE; PS50172; BRCT; 1.
KM Nuclear protein; Coiled coil.
FT DOMAIN 269 329 COILED COIL (POTENTIAL).
FT DOMAIN 323 416 BRCT.
FT DOMAIN 453 543 COILED COIL (POTENTIAL).
FT DOMAIN 537 557 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 455 470 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 472 483 ASP/GLU-RICH (ACIDIC).
SQ SEQUENCE 582 AA; 67770 MW; 51ED59F83DDEBF34 CRC64;

Query Match 27.1%; Score 58.5; DB 1; Length 582;
Best Local Similarity 34.5%; Pred. No. 12;
Matches 20; Conservative 6; Mismatches 11; Indels 21; Gaps 3;

OY 1 EDDERSTDSQOCSSDEDFEETA-----QVSPRGKEKQWRAR 41
Db 466 EDDDEDEDDQ--SEDEBAEENLAEMERKSGCKSLSVKTP--GAKAKENRAR 519

RESULT 12
IE63_HSV1E STANDARD; PRT; 511 AA.
AC P36295;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)

```

DE Transcriptional regulator IE63 (VMM63) (ICP27).
GN UL54.
OS Herpes simplex virus (type 1 / strain HFEM).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10303;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93070559; PubMed=1332274;
RA Rosen-Wolf A., Frank S., Raab K., Moyal M., Becker Y., Darai G.;
RT "Determination of the coding capacity of the BamHI DNA fragment B of
RT apathogenic Herpes simplex virus type 1 strain HFEM by DNA nucleotide
RT sequence analysis."
RL Virus Res. 25:189-199(1992).
CC -1- FUNCTION: INVOLVED IN THE DOWN-REGULATION OF VIRAL IMMEDIATE-EARLY
CC GENES AND ACTS IN COMBINATION WITH ICP0 AND ICP4 AS AN ACTIVATOR
CC OF LATE GENES.
CC -1- PTM: PHOSPHORYLATED.
CC -1- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL54,
CC HSV-2 UL54, EBV-1 5, VZV 4, EBV BMLF1, HCMV UL69, AND HVS-1 57.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M90438; -; NOT ANNOTATED_CDS.
CC PIR; A48560; A48560.
CC DR Early protein; Transcription regulation; Activator; DNA-binding;
CC KW Phosphorylation.
CC SEQUENCE 511 AA; 55142 MW; 7BE7A8F841A98174 CRC64;
SQ
Query Match 26.9%; Score 58; DB 1; Length 511;
Best Local Similarity 40.7%; Pred. No. 12;
Matches 11; Conservative 6; Mismatches 10; Indels 0; Gaps 0;
OY 3 DERSTDSSQCCSSEDDIFFEETAGVSP 29
Db 33 DLESDDSSGECSSSDDEMDPHGEDP 59

```

```

CC -1- FUNCTION: INVOLVED IN THE DOWN-REGULATION OF VIRAL IMMEDIATE-EARLY
CC GENES AND ACTS IN COMBINATION WITH ICP0 AND ICP4 AS AN ACTIVATOR
CC OF LATE GENES.
CC -1- PTM: PHOSPHORYLATED.
CC -1- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL54,
CC HSV-2 UL54, EBV-1 5, VZV 4, EBV BMLF1, HCMV UL69, AND HVS-1 57.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X14112; CA32290.1; -.
CC DR PIR; I30089; WMBEY4.
CC DR Early protein; Transcription regulation; Activator; DNA-binding;
CC KW Phosphorylation.
CC SEQUENCE 512 AA; 55252 MW; 97DF74A2B7E63A85 CRC64;
SQ
Query Match 26.9%; Score 58; DB 1; Length 512;
Best Local Similarity 40.7%; Pred. No. 12;
Matches 11; Conservative 6; Mismatches 10; Indels 0; Gaps 0;
OY 3 DERSTDSSQCCSSEDDIFFEETAGVSP 29
Db 33 DLESDDSSGECSSSDDEMDPHGEDP 59

```

```

RESULT 13
ID IE63_HSV11 STANDARD; PRT; 512 AA.
AC P10238;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Transcriptional regulator IE63 (VMM63) (ICP27).
GN UL54.
OS Herpes simplex virus (type 1 / strain 17).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88274327; PubMed=2839594;
RA McGeoch D.J., Dalrymple M.A., Davison A.J., Dolan A., Frame M.C.,
RA McNab D., Perry L.J., Scott J.B., Taylor P.;
RT "The complete DNA sequence of the long unique region in the genome of
RT herpes simplex virus type 1."
RL J. Gen. Virol. 69:1531-1574(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89036163; PubMed=2846760;
RA Perry L.J., McGeoch D.J.;
RT "The DNA sequences of the long repeat region and adjoining parts of
RT the long unique region in the genome of herpes simplex virus type
RT 1."
RL J. Gen. Virol. 69:2831-2846(1988).

```

```

RESULT 14
ID ABAA_EMENT STANDARD; PRT; 796 AA.
AC P20945;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Regulatory protein abaa.
GN ABAA.
OS Emmericella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; Emmericella.
OX NCBI_TaxID=162425;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FGSC 26;
RX MEDLINE=89249350; PubMed=2655931;
RA Miralbero P.M., Adams T.H., Timberlake W.E.;
RT "Interactions of three sequentially expressed genes control temporal
RT and spatial specificity in Aspergillus development."
RL Cell 57:859-868(1989).
RN [2]
RP TEA DOMAIN.
RX MEDLINE=91300541; PubMed=2070413;
RA Buergelein T.R.;
RT "The TEA domain: a novel, highly conserved DNA-binding motif."
RL Cell 66:11-12(1991).
CC -1- FUNCTION: CONTROLS TEMPORAL AND SPATIAL SPECIFICITY IN ASPERGILLUS
CC DEVELOPMENT. EXPRESSION OF ABAA LEADS TO ACTIVATION OF BRLA AND
CC WETA, CESSATION OF VEGETATIVE GROWTH, AND ACCELERATED CELLULAR
CC VACUOLIZATION.
CC -1- FUNCTION: BRLA, ABAA & WETA ARE PIVOTAL REGULATORS OF CONIDIOPHORE
CC DEVELOPMENT AND CONIDIUM MATURATION. THEY ACT INDIVIDUALLY AND
CC TOGETHER TO REGULATE THEIR OWN EXPRESSION AND THAT OF NUMEROUS
CC OTHER SPORULATION-LOCATIONS.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: Contains 1 TEA DNA-binding domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: J04850; AAA33286.1; -

DR TRANSPAC; T01085; -

DR InterPro; IPR000818; TEA/ATTSdom.

DR Pfam; PF01285; TEA; 1.

DR PRINTS; PRO0065; TEADOMAIN.

DR SMART; SM00426; TEA; 1.

DR PROSITE; PS00554; TEA_DOMAIN; 1.

DR Developmental protein; Conidiation; Transcription regulation;

KM Activator; DNA-binding; Nuclear protein.

FT DNA_BIND 135 TEA-DOMAIN.

FT DOMAIN 341 362 LEUCINE-ZIPPER LIKE.

SQ SEQUENCE 796 AA; 89170 MW; 79429E484F56E035 CRC64;

Query Match 26.9%; Score 58; DB 1; Length 796;
 Best Local Similarity 35.3%; Pred. No. 20;
 Matches 12; Conservative 6; Mismatches 14; Indels 2; Gaps 1;

OY 8 DSSQCCSSEDEDIFETAOVSPRGKEXRQ 41
 Db 135 DGEFWSDELEDAFQALANPPMG--RRKMSR 166

RESULT 15

YCFL EPIVI

ID _YCFL EPIVI STANDARD; PRT; 1738 AA.

AC 000383;

DT 01-APR-1993 (Rel. 25, Created)

DT 01-APR-1993 (Rel. 25, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Hypothetical 208 kDa protein ycf1 (ORF 1738).

GN YCF1.

OS Epifagus virginiana (Beechdrops).

OG Chloroplast.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; lamids; Lamiales; Orobanchaceae; Orobanchaceae; Epifagus.

OX NCBI_TaxID=4177;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=92114084; PubMed=1731088;
 RA Wolfe K.H., Morden C.W., Palmer J.D.;
 RT "Small single-copy region of plastid DNA in the non-photosynthetic
 RT angiosperm Epifagus virginiana contains only two genes. Differences
 RT among dicots, monocots and bryophytes in gene organization at a non-
 RT bioenergetic locus.";
 RT J. Mol. Biol. 223:95-104(1992).

RL [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=93066301; PubMed=1332054;
 RA Wolfe K.H., Morden C.W., Palmer J.D.;
 RT "Function and evolution of a minimal plastid genome from a
 RT nonphotosynthetic parasitic plant.";
 RT Proc. Natl. Acad. Sci. U.S.A. 89:10648-10652(1992).

RL [1]

CC -1- FUNCTION: NOT YET KNOWN.

CC -1- SIMILARITY: BELONGS TO THE YCF1 FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; X61368; CAA43644.1; -

DR EMBL; M81884; AAA65870.1; -

DR PIR; S20614; S20614

KM Chloroplast; Hypothetical protein.

SQ SEQUENCE 1738 AA; 208318 MW; 65C63F63BC8364B CRC64;

Query Match 26.9%; Score 58; DB 1; Length 1738;
 Best Local Similarity 48.4%; Pred. No. 48;
 Matches 15; Conservative 2; Mismatches 12; Indels 2; Gaps 1;

OY 7 TDSQCCSSEDEDIFETAOVSPRGKEXRQ 37
 Db 245 TLKKEQOKVDEDI--ETVQMPLGLEKR 273

Search completed: July 24, 2003, 11:54:21
 Job time : 5.86091 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 11:47:53 ; Search time 9.12445 Seconds
(without alignment)
1159.539 Million cell updates/sec

Title: US-09-991-681-28

Perfect score: 216
Sequence: 1 EDDERSTDSQQCSSEDEDI.....ETTAQVSPRGKEKRWRR 41

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_23:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phage:*

10: sp_plant:*

11: sp_prodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_virus:*

16: sp_bacteriopl:*

17: sp_archaeopl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	216	100.0	1770	4 Q9ULH6	Q9ULH6 homo sapien
2	195.5	90.5	589	4 Q8N4Y4	Q8N4Y4 homo sapien
3	195.5	90.5	592	4 Q96CH9	Q96CH9 homo sapien
4	66	30.6	584	11 Q9CR92	Q9CR92 mus musculu
5	65.5	30.3	521	4 Q8N5Y2	Q8N5Y2 homo sapien
6	65.5	30.3	521	4 Q9Y5Z8	Q9Y5Z8 homo sapien
7	64	29.6	687	4 Q8NB06	Q8NB06 homo sapien
8	63.5	29.4	1114	5 Q9NED2	Q9NED2 leishmania
9	63	29.2	707	11 Q99K50	Q99K50 mus musculu
10	62.5	28.9	642	5 Q9V7N4	Q9V7N4 drosophila
11	62.5	28.9	652	5 Q8T092	Q8T092 drosophila
12	62	28.7	422	4 Q8WUN3	Q8WUN3 homo sapien
13	61	28.2	1347	10 Q38969	Q38969 arabidopsis
14	61	28.2	1854	10 Q9C606	Q9C606 arabidopsis
15	61	28.2	2599	10 Q9C727	Q9C727 arabidopsis
16	60.5	28.0	715	11 Q9QZX1	Q9QZX1 ratus norv

17	60	27.8	450	5 Q8MYL8	Q8MYL8 caenorhabdi
18	60	27.8	484	11 Q9CS98	Q9CS98 mus musculu
19	60	27.8	1144	5 Q9NH77	Q9NH77 drosophila
20	60	27.8	1497	4 Q8IWB6	Q8IWB6 homo sapien
21	60	27.8	1657	10 Q8W0X4	Q8W0X4 zea mays (m
22	59.5	27.5	292	10 Q9LVH1	Q9LVH1 arabidopsis
23	59.5	27.5	413	4 Q96AJ1	Q96AJ1 homo sapien
24	59.5	27.5	1142	5 Q95XD2	Q95XD2 caenorhabdi
25	59.5	27.5	2287	5 Q8IS20	Q8IS20 plasmodium
26	59	27.3	498	5 Q8MTN8	Q8MTN8 trichinella
27	59	27.3	550	11 Q8B0D8	Q8B0D8 mus musculu
28	59	27.3	586	16 Q9H2D4	Q9H2D4 pseudomonas
29	59	27.3	707	11 Q8CE30	Q8CE30 mus musculu
30	59	27.3	957	4 Q9BX79	Q9BX79 homo sapien
31	59	27.3	1348	4 Q8ND97	Q8ND97 homo sapien
32	58.5	27.1	182	4 Q9NSU7	Q9NSU7 homo sapien
33	58.5	27.1	517	4 Q96LT9	Q96LT9 homo sapien
34	58.5	27.1	541	4 Q96U18	Q96U18 homo sapien
35	58.5	27.1	566	4 Q9UP68	Q9UP68 homo sapien
36	58	26.9	1781	4 Q9UKX0	Q9UKX0 homo sapien
37	58	26.9	1784	6 Q8MM13	Q8MM13 macaca fasc
38	58	26.9	1890	4 Q9UKW3	Q9UKW3 macaca fasc
39	58	26.9	2072	4 Q8RYB5	Q8RYB5 homo sapien
40	58	26.9	2073	4 Q9UKW2	Q9UKW2 homo sapien
41	58	26.9	3328	5 Q8IUG6	Q8IUG6 plasmodium
42	58	26.9	4065	3 Q9P4Z1	Q9P4Z1 neurospora
43	57.5	26.6	88	16 Q93IC4	Q93IC4 rhizobium m
44	57.5	26.6	1334	11 Q8CAP0	Q8CAP0 mus musculu
45	57	26.4	235	13 Q8AVR0	Q8AVR0 xenopus lae

ALIGNMENTS

RESULT 1

ID Q9ULH6 PRELIMINARY; PRT; 1770 AA.

AC Q9ULH6; Q96P46;

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)

DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)

DE B1G3.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Hong W.;

RT "K1A1244 as a novel distantly related member (B1G3) of the B1G1/Sec7p subfamily of ARF GTPs.";

RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF413080; AL04174.1; -.

DR InterPro; IPR00904; Sec7.

DR SMART; SM00222; Sec7; 1.

KW Hypothetical protein.

SQ SEQUENCE 1770 AA; 195845 MW; 5E996E36A6F2AB4 CRC64;

Query Match 100.0%; Score 216; DB 4; Length 1770;
Best Local Similarity 100.0%; Pred. No. 4.5e-20;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EDDERSTDSQQCSSEDEDIFEETPAQVSPRGKEKRWRR 41
Db 1436 EDDERSTDSQQCSSEDEDIFEETPAQVSPRGKEKRWRR 1476

RESULT 2

ID Q8N4Y4 PRELIMINARY; PRT; 589 AA.

AC Q8N4Y4;

DT 01-OCT-2002 (TREMBLrel. 22, Created)

DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)

```

DT 01-OCT-2002 (TREMBlrel. 22, last annotation update)
DE Similar to KIAA1244 protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC033191; AAH33191.1; -.
FT NON TER
SQ SEQUENCE 589 AA; 66086 MW; 8041EA348DE65F7 CRC64;

Query Match
Best Local Similarity 90.5%; Score 195.5; DB 4; Length 589;
Matches 39; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 EDDERSTDSQCCSDEDFEETAAVSPPRGKERQWRA 40
DB 256 EDDERSTDSQCCSDEDFEETAAVSPPRGK-KRQWRA 294

RESULT 3
Q96CH9 PRELIMINARY; PRT; 592 AA.
AC Q96CH9;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC014227; AAH14227.1; -.
KW Hypothetical protein.
FT NON TER
SQ SEQUENCE 592 AA; 66400 MW; F4A1E807B0DF47B5 CRC64;

Query Match
Best Local Similarity 90.5%; Score 195.5; DB 4; Length 592;
Matches 39; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 EDDERSTDSQCCSDEDFEETAAVSPPRGKERQWRA 40
DB 259 EDDERSTDSQCCSDEDFEETAAVSPPRGK-KRQWRA 297

RESULT 4
Q9CR92 PRELIMINARY; PRT; 584 AA.
AC Q9CR92;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, last annotation update)
DE 4921513E08rik protein (Hypothetical aspartic acid-rich region/glutamic
DE acid-rich region containing protein).
GN 4921513E08RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=21085660; PubMed=11217851;

```

```

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochava H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirml L.W., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momberees P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilting L.,
RA Wymshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851;
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK015646; BAB29914.1; -.
DR EMBL; AK014878; BAB29599.1; -.
DR EMBL; AK030023; BAC26742.1; -.
DR MGI; MGI:1913967; 4921513E08rik.
KW Hypothetical protein.
SQ SEQUENCE 584 AA; 67255 MW; 0F360DF47AD78E18 CRC64;

Query Match
Best Local Similarity 30.6%; Score 66; DB 11; Length 584;
Matches 14; Conservative 7; Mismatches 18; Indels 0; Gaps 0;

QY 1 EDDERSTDSQCCSDEDFEETAAVSPPRGKERQWRA 39
DB 125 DEDEDEDEDEDEDEDEDEDEDEDEGERDRPEKRGKREKR 163

RESULT 5
Q8N5Y2 PRELIMINARY; PRT; 521 AA.
AC Q8N5Y2;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, last annotation update)
DE Male-specific lethal 3-like 1 (Drosophila).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC031210; AAH31210.1; -.
DR Genew; HGNC:7370; MSL3L1.
DR InterPro; IPR000953; Chromo.
DR SMART; SMO0298; CHROMO; 1.
SQ SEQUENCE 521 AA; 59823 MW; 6DFFB9BE183DDCFB9 CRC64;

Query Match
Best Local Similarity 30.3%; Score 65.5; DB 4; Length 521;
Matches 17; Conservative 6; Mismatches 18; Indels 7; Gaps 1;

QY 1 EDDERSTDSQCCS-----SDEBDIFEETAAVSPPRGKERQWRA 41

```

Db 123 ENDENSLSSSSDCSENKDEISEESDIEKTEVKEPELQTRREMEER 170

RESULT 6

Q9Y5Z8 PRELIMINARY; PRT; 521 AA.
AC Q9Y5Z8;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Male-specific lcthal-3 homolog 1.
GN MSJ3L1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99326525; PubMed=10395802;
RA Prakash S.K., Van den Veyver I.B., Franco B., Volta M., Ballabio A.,
RA Zoghbi H.Y.;
RT "Characterization of a novel chromo domain gene in Xp22.3 with
RT homology to Drosophila msl-3."
RL Genomics 59:77-84(1999).
DR EMBL; AF117065; AAD38499.1; -
DR InterPro; IPR000953; Chromo.
DR SMART; SMO0298; CHROMO.1.
SQ SEQUENCE 521 AA; 59807 MW; 0C65A40FD874CD CRC64;

Query Match 30.3%; Score 65.5; DB 4; Length 521;
Best Local Similarity 35.4%; Pred. No. 2;
Matches 17; Conservative 6; Mismatches 18; Indels 7; Gaps 1;

Qy 1 EDDERSTSSQCCS-----SEDDIFEEFTAQVSPPGKEXKQWRAR 41
Db 123 ENDENSLSSSSDCSENKDEISEESDIEKTEVKEPELQTRREMEER 170

RESULT 7

Q8NB06 PRELIMINARY; PRT; 687 AA.
AC Q8NB06;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hypothetical protein FLJ34423.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Nihi T., Ota T., Nakagawa S., Senoh A., Mizuguchi H., Inagaki H.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
RA Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Yamaehita H., Matsuo K., Nakamura Y., Sekine M.,
RA Kituchi H., Kanda K., Wagaatsuma M., Murakawa K., Kanehori K.,
RA Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B., Suzuki Y.,
RA Sugano S., Nagahari K., Maehuo Y., Nagai K., Isogai T.;
RT "NEDO human cDNA sequencing project."
RL Submitted (JUL-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; AK091742; BAC03738.1; -
DR InterPro; IPR00504; RNA_rec_mot.
DR Pfam; PF00076; rrm; 4.
DR SMART; SMO0360; RRM; 4.
DR PROSITE; PS50102; RRM; 4.
DR PROSITE; PS00030; RRM_RNP_1; 3.
KW Hypothetical protein.
SQ SEQUENCE 687 AA; 74324 MW; 3494C6825B54A32D CRC64;

Query Match 29.6%; Score 64; DB 4; Length 687;
Best Local Similarity 35.3%; Pred. No. 4.2;

Matches 12; Conservative 6; Mismatches 16; Indels 0; Gaps 0;
Qy 1 EDDERSTSSQCCSSEDDIFEEFTAQVSPPGKEXKQWRAR 41
Db 164 EDEDEDEDEDEDEDEDEDEDESEEMETTPAKKK 197

RESULT 8

Q9NED2 PRELIMINARY; PRT; 1114 AA.
AC Q9NED2;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Hypothetical 119.2 kDa protein.
GN p1105.11.
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RA Ivens A.C., Murphy L., Quail M., Harris D., Oliver K., Lawson D.,
RA Rajandream M.A., Barrell B.G.;
RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RX MEDLINE=98146435; PubMed=9477341;
RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
RA Smith D.F.;
RT "A physical gap of the Leishmania major Friedlin genome."
RL Genome Res. 8:135-145(1998).
DR EMBL; AL139794; CAB75569.1; -
KW Hypothetical protein.
SQ SEQUENCE 1114 AA; 119250 MW; DB8A995FA5793A17 CRC64;

Query Match 29.4%; Score 63.5; DB 5; Length 1114;
Best Local Similarity 28.3%; Pred. No. 8.2;
Matches 13; Conservative 11; Mismatches 15; Indels 7; Gaps 1;

Qy 3 DERSTSSQCCSSEDDIFEEFTAQVSPPGKEXKQWRAR 41
Db 63 EEDVVSATTCASPEAVLEQPRVAPDQORPGRORAPRRRMSR 108

RESULT 9

Q99K50 PRELIMINARY; PRT; 707 AA.
AC Q99K50;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Nucleolin.
GN NCL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strauberg R.;
RL Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; BC005460; AAH05460.1; -
DR HSSP; P09651; IHAI.
DR MGD; MGI:97286; NCI.
DR InterPro; IPR00504; RNA_rec_mot.
DR Pfam; PF00076; rrm; 4.
DR SMART; SMO0360; RRM; 4.
DR PROSITE; PS50102; RRM; 4.
DR PROSITE; PS00030; RRM_RNP_1; 3.
SQ SEQUENCE 707 AA; 76779 MW; 44B42C6E146E5D50 CRC64;

RA Dunn P., Ekgu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizer L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 RA Kim C.U., Koo H.L., Kremetskaya I., Kurtz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maiti R., Marziani A.,
 RA Miltischer J., Miranda M., Nguyen M., Nieman W.C., Osborne B.I.,
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzer S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
 RA Utecherback T., Van Aken S., Vayenberg M., Vysotskaya V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.,
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 thaliana."
 RL Nature 408:816-820(2000).
 DR EMBL; AC073943; AAG50951.1; -.
 KM Hypothetical protein.
 SQ SEQUENCE 2539 AA; 292445 MW; 6FAF3CA592F4A3C CRC64;

Query Match 28.2%; Score 61; DB 10; Length 2539;
 Best Local Similarity 57.1%; Pred. No. 44;
 Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 9 SSQCCSSEDDIFFEETAVSP 29
 |||:||||:|:|
 Db 1968 SSLQCTEDDEVDVEEADRYVP 1988

Search completed: July 24, 2003, 11:56:59
 Job time : 12.1245 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: July 24, 2003, 11:58:19 ; Search time 131.965 Seconds

(without alignments)
623.047 Million cell updates/sec

Title: US-09-991-681-27

Perfect score: 518
Sequence: 1 RIRAMQVFMPLDTQCSPKT.....VRQAVREMLGRVGRVDIIV 518

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1107863 seqs, 158726573 residues

Word size: 0

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

A_Geneseq_19Jun03:*

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	518	100.0	518	19 AAW85472	PS118 protein enco
2	518	100.0	518	23 AAO19165	Human prostate-spe
3	518	100.0	518	23 AAM50809	PS118 prostate mar
4	518	100.0	1807	22 ABG09728	Novel human diagno
5	518	100.0	1839	23 ABP64835	Human protein SGO
6	451	87.1	1982	22 ABG09731	Novel human diagno
7	192	37.1	192	20 AAY48248	Human prostate can
8	180	34.7	180	20 AAY48216	Human prostate can
9	49	9.5	49	23 AAO19169	Human prostate-spe

10	49	9.5	49	23 AAM50813	PS118 prostate mar
11	43	8.3	50	22 ABG09729	Novel human diagno
12	41	7.9	41	23 AAO19166	Human prostate-spe
13	41	7.9	41	23 AAM50810	PS118 prostate mar
14	40	7.7	40	23 AAO19168	Human prostate-spe
15	40	7.7	40	23 AAM50812	PS118 prostate mar
16	35	6.8	35	23 AAO19167	Human prostate-spe
17	35	6.8	35	23 AAM50811	PS118 prostate mar
18	9	1.7	216	24 ABP81377	Human prostate mar
19	9	1.7	216	24 ABU01364	S. pneumoniae type
20	8	1.5	32	22 ABG49938	Human liver peptid
21	8	1.5	32	22 ABB29911	Peptide #2562 enco
22	8	1.5	32	22 ABB35089	Peptide #2595 enco
23	8	1.5	32	22 ABB20507	Protein #2506 enco
24	8	1.5	32	22 AAM55910	Human brain expres
25	8	1.5	32	22 AAM68280	Human bone marrow
26	8	1.5	32	22 AAM16103	Peptide #2537 enco
27	8	1.5	32	22 AAM28594	Peptide #2631 enco
28	8	1.5	32	22 AAM03830	Peptide #2512 enco
29	8	1.5	32	23 ABG37821	Human peptid enco
30	8	1.5	121	22 AAB95737	Human protein sequ
31	8	1.5	343	21 AAG20817	Arabidopsis thalia
32	8	1.5	351	23 AAE18955	Human cell cycle p
33	8	1.5	427	21 AAG20816	Arabidopsis thalia
34	8	1.5	448	21 AAG20815	Arabidopsis thalia
35	8	1.5	477	15 AAR43999	Human OTC. Homo s
36	8	1.5	483	23 ABB47760	Listeria monocytog
37	8	1.5	578	12 AAR10098	Alpha subunit of A
38	8	1.5	578	12 AAR10098	Enterococcus faeca
39	8	1.5	1031	22 AAG35339	Human complementar
40	7	1.4	10	22 AAG36976	Peptide from human
41	7	1.4	13	20 AAW86038	Peptide from human
42	7	1.4	13	20 AAW86065	Peptide from human
43	7	1.4	13	20 AAW86066	Peptide from human
44	7	1.4	13	20 AAW86072	Peptide from human
45	7	1.4	13	20 AAW86046	Peptide from human
					Peptide from human

ALIGNMENTS

RESULT 1	
AAW85472	
ID AAW85472 standard; Protein: 518 AA.	
XX	
AC AAW85472;	
XX	
DT 25-MAR-2003 (updated)	
DT 25-FEB-1999 (first entry)	
XX	
DE PS118 protein encoded by consensus sequence.	
XX	
KM EST clone; PS118; prostate tumour tissue; prostatic disease; cancer.	
XX	
OS Homo sapiens.	
XX	
PN MO9848054-A1.	
XX	
PD 29-OCT-1998.	
XX	
PF 23-APR-1998; 98WC-US08239.	
XX	
PR 23-APR-1997; 97US-0842385.	
XX	
PA (ABBO) ABBOTT LAB.	
XX	
PI Billing-medel PA, Cohen M, Colpites TL, Friedman PN, Gordon J;	
PI Granados EN, Hodges SC, Klass MR, Kratochvil JD, Robertsrapp L;	
PI Russell JC, Stroupe SD;	
XX	
DR WPI, 1998-610000/51.	
DR N-PSDB; AAW82812.	
XX	

PT New P118 nucleic acid and proteins - used for diagnosis and
PT treatment of prostatic disease, especially cancer, and also for drug
PT screening

Claim 17; Page 93-94; 117pp; English.

CC The present sequence is encoded by consensus P118 sequence derived from
CC overlapping EST (expressed sequence tag) clones AAV82803-10. The clones
CC were identified from a cDNA library made from prostate tumour tissue.
CC Recombinant P118 protein is used to detect P118-specific antibodies,
CC to raise antibodies for detection of P118 antigens, to screen for
CC specific binding agents (potential therapeutics, and to isolate specific
CC antibodies from serum. Detection of P118 protein or nucleic acid, which
CC are prostate related, and altered or elevated in prostatic disease, is
CC used for detection, diagnosis, staging, monitoring and prognosis of
CC prostatic disease, particularly cancer, and to identify subjects at
CC risk.
CC (Updated on 25-MAR-2003 to correct PI field.)

CC Sequence 518 AA;

Query Match 100.0%; Score 518; DB 19; Length 518;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIRAMAQQVFMILDTQCSPTNNFPHQSCOLIIELPDEKPNHTKKSVSFREIVSL 60
DB 1 RIRAMAQQVFMILDTQCSPTNNFPHQSCOLIIELPDEKPNHTKKSVSFREIVSL 60
QY 61 SHOVLQNLVYDILIEEFVKGSPGEEKTIQVPEAKLAGFLRYISQNLAVIFDLIDS 120
DB 61 SHOVLQNLVYDILIEEFVKGSPGEEKTIQVPEAKLAGFLRYISQNLAVIFDLIDS 120
QY 121 TAREPDTSPGLKCLKLVKVSIGGANLYROSAMSFNIYFHALVCAVLTNOETITAEQV 180
DB 121 TAREPDTSPGLKCLKLVKVSIGGANLYROSAMSFNIYFHALVCAVLTNOETITAEQV 180
QY 181 VLFEDDERSTDSQOCSEDEDIFFETQAVSPRGKEKQVAMPPLISVQVSNADW 240
DB 181 VLFEDDERSTDSQOCSEDEDIFFETQAVSPRGKEKQVAMPPLISVQVSNADW 240
QY 241 LVKRLHKLKCMELCNVYIQWHLLENQMEBPPIFKGDPFPIIPSPQSESTPTGFS 300
DB 241 LVKRLHKLKCMELCNVYIQWHLLENQMEBPPIFKGDPFPIIPSPQSESTPTGFS 300
QY 301 TPSEDDRSQSRHEHGESISLAKAGGDDLPPSPVYKPKDPSRKKEWENAGNKI 360
DB 301 TPSEDDRSQSRHEHGESISLAKAGGDDLPPSPVYKPKDPSRKKEWENAGNKI 360
QY 361 KTISTLMTEYKRRKQOHNLSAPFKVYKKEKGPFGRGQDPSPIORPHLMDQOM 420
DB 361 KTISTLMTEYKRRKQOHNLSAPFKVYKKEKGPFGRGQDPSPIORPHLMDQOM 420
QY 421 FSAGBELLRQDKRPPSGTSSLSVSRDABAQIOATMNVLTVALNOIILPDQFT 480
DB 421 FSAGBELLRQDKRPPSGTSSLSVSRDABAQIOATMNVLTVALNOIILPDQFT 480
QY 481 PAVFPCISQLTCHVTDIRQAVREWLGRVGRVYDIIV 518
DB 481 PAVFPCISQLTCHVTDIRQAVREWLGRVGRVYDIIV 518

RESULT 2
ID AAO19165 standard; Protein; 518 AA.
AC AAO19165;
XX 27-NOV-2002 (first entry)
DB Human prostate-specific P118 protein fragment #1.
XX Human; prostate; prostate-specific sequence; prostate cancer; P118;

KW EST; expressed sequence tag; cytostatic; gene therapy.

OS Homo sapiens.

PN US2002086316-A1.

PD 04-JUL-2002.

PF 26-NOV-2001; 2001US-0991681.

PR 23-APR-1998; 98US-0065383.

PR 23-APR-1997; 97US-0842385.

XX (BILL/) BILLINGEL P A.

PA (COHE/) COHEN M.

PA (COLP/) COLPITTS T L.

PA (FRIE/) FRIEDMAN P N.

PA (GORD/) GORDAN J.

PA (GRAN/) GRANADOS E N.

PA (HODG/) HODGES S C.

PA (KLAS/) KLAS M R.

PA (KRAT/) KRATOCHVIL J D.

PA (ROBE/) ROBERTS-RAPP L.

PA (RUS/) RUSSELL J C.

PA (STRO/) STROUPE S D.

XX Billingsel PA, Cohen M, Colpitts TL, Friedman PN, Gordan J,

PI Granados EN, Hodges SC, Klaas MR, Kratochvil JD, Roberts-Rapp L,

PI Russell JC, Stroupe SD;

DR WPI, 2002-665429/71.

PT Novel P118 polypeptide for detecting, diagnosing, staging, monitoring,

PT prognosticating, preventing, treating, or determining predisposition of

PT individual to diseases and conditions of prostate, e.g. prostate

PT cancer -

XX Claim 17; Page 42-43; 58pp; English.

PS The present invention relates to a number of prostate-specific sequences

CC derived from the human P118 gene. These can be used in the detection,

CC monitoring and treatment of prostate diseases, particularly prostate

CC cancer. The present sequence is a P118 protein fragment of the

CC invention. The coding sequences of the invention were isolated from a

CC prostate tissue expressed sequence tag (EST) library.

CC Sequence 518 AA;

QY Query Match 100.0%; Score 518; DB 23; Length 518;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIRAMAQQVFMILDTQCSPTNNFPHQSCOLIIELPDEKPNHTKKSVSFREIVSL 60
DB 1 RIRAMAQQVFMILDTQCSPTNNFPHQSCOLIIELPDEKPNHTKKSVSFREIVSL 60
QY 61 SHOVLQNLVYDILIEEFVKGSPGEEKTIQVPEAKLAGFLRYISQNLAVIFDLIDS 120
DB 61 SHOVLQNLVYDILIEEFVKGSPGEEKTIQVPEAKLAGFLRYISQNLAVIFDLIDS 120
QY 121 TAREPDTSPGLKCLKLVKVSIGGANLYROSAMSFNIYFHALVCAVLTNOETITAEQV 180
DB 121 TAREPDTSPGLKCLKLVKVSIGGANLYROSAMSFNIYFHALVCAVLTNOETITAEQV 180
QY 181 VLFEDDERSTDSQOCSEDEDIFFETQAVSPRGKEKQVAMPPLISVQVSNADW 240
DB 181 VLFEDDERSTDSQOCSEDEDIFFETQAVSPRGKEKQVAMPPLISVQVSNADW 240
QY 241 LVKRLHKLKCMELCNVYIQWHLLENQMEBPPIFKGDPFPIIPSPQSESTPTGFS 300
DB 241 LVKRLHKLKCMELCNVYIQWHLLENQMEBPPIFKGDPFPIIPSPQSESTPTGFS 300
QY 301 TPSEDDRSQSRHEHGESISLAKAGGDDLPPSPVYKPKDPSRKKEWENAGNKI 360

```

Db 301 TPSEDRSQSREHNGESLTKAGGDDLPPSPVKEKDPKSKKEMWENGNKIYTWAD 360
Qy 361 KTISKMTTEYKKRQOHNLSAFPKEVVEKKGEPGPGDPSLQRPQHLMDQGQRHS 420
Db 361 KTISKMTTEYKKRQOHNLSAFPKEVVEKKGEPGPGDPSLQRPQHLMDQGQRHS 420
Qy 421 FSAGPELLRQDKRPRSGSTGSSLSVSVRDAEAOIQAWTNVLTVLNOIQILPDQFTALQ 480
Db 421 FSAGPELLRQDKRPRSGSTGSSLSVSVRDAEAOIQAWTNVLTVLNOIQILPDQFTALQ 480
Qy 481 PAVFPCTISQLTCHVTDIRVQAVREMLGRVGRVYDIIV 518
Db 481 PAVFPCTISQLTCHVTDIRVQAVREMLGRVGRVYDIIV 518

RESULT 3
AAM50809
ID AAM50809 standard; Protein; 518 AA.
AC AAM50809;
XX 01-MAY-2002 (first entry)
XX PS118 prostate marker partial sequence.
XX PS118 prostate marker: prostate cancer; tumour; metastasis;
XX benign prostatic hyperplasia; prostatic intraepithelial neoplasia;
XX prostaticitis; human; diagnosis; therapy; vaccine.
XX Homo sapiens.
XX OS
XX US2001055758-A1.
XX 27-DEC-2001.
XX 23-APR-1998; 98US-0065383.
XX 23-APR-1997; 97US-0842385.
XX PA (BILL/) BILTING-MEDEL P A.
XX PA (COHE/) COHEN M.
XX PA (COPL/) COPLITTIS T L.
XX PA (FRIE/) FRIEDMAN P N.
XX PA (GORD/) GORDON J.
XX PA (GRAN/) GRANADOS E N.
XX PA (HODG/) HODGES S C.
XX PA (KLAS/) KLAS M R.
XX PA (KRAT/) KRATOCHVIL J D.
XX PA (ROBE/) ROBERTS-RAPP L.
XX PI Billing-Medel PA, Cohen M, Coplittis TL, Friedman PN, Gordon J,
PI Granados EN, Hodges SC, KLAS MR, Kratochvil JD, Roberts-Rapp L;
PI Russell JC, Stroupe SD;
XX DR MPI: 2002-187683/24.
XX N-PSDB; ABA91651.
XX PT Detecting presence of target PS118 polynucleotide in test sample,
XX useful for detecting, diagnosing, staging, monitoring, prognosticating,
XX preventing or treating or determining predisposition to prostate
XX disease -
XX PS Claim 17; Page 42-43; 57pp; English.
XX CC The present sequence is that of a human prostate-specific PS118
XX polypeptide, as predicted from a partial consensus cDNA sequence
XX (see ABA91651), and lacking the N-terminal region. The PS118
XX consensus sequence is found at least 12 times more often in
XX prostate than in non-prostate tissue. PS118 polypeptides,
XX including derivatives of the present sequence, polynucleotides,
XX antibodies, agonists and inhibitors are useful for detecting
XX diagnosing, staging, monitoring, prognosticating, preventing and

```

```

CC treating (including by genetic immunisation), or determining the
CC predisposition of an individual to diseases and conditions of the
CC prostate, such as benign prostatic hyperplasia, prostaticitis,
CC prostatic intraepithelial neoplasia, prostate cancer, tumours and
CC metastases. The PS118 polypeptides can be produced by expression
CC of PS118 polynucleotides in transfected host cells. The methods
CC and reagents of the invention may provide an early means of
CC detecting diseases of the prostate and may also provide new markers
CC which can differentiate between the clinically important and
CC unimportant prostate cancers without the use of surgery.
XX SQ Sequence 518 AA;
XX Query Match 100.0%; Score 518; DB 23; Length 518;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 RIRAMAQOVFMDLQCSPTNNPFDHAQSCQIILBPPDEKPNQHTKKSVPREIVSLL 60
Db 1 RIRAMAQOVFMDLQCSPTNNPFDHAQSCQIILBPPDEKPNQHTKKSVPREIVSLL 60
Qy 61 SHOVLQNLVYDILLEFVVKSPGEEKTIQVPAKLAGFIRYISQNLAVIFPLLDYSR 120
Db 61 SHOVLQNLVYDILLEFVVKSPGEEKTIQVPAKLAGFIRYISQNLAVIFPLLDYSR 120
Qy 121 TAREPDTSPGLKCLLKXVSGIGGANLYROSASFNIYFALVCAVLTNOETITAEQVK 180
Db 121 TAREPDTSPGLKCLLKXVSGIGGANLYROSASFNIYFALVCAVLTNOETITAEQVK 180
Qy 181 VLFEDDERSTDSQCSSEDEDIFFETAOVSPRGKEKQWRAMPPLSVQVSNADWW 240
Db 181 VLFEDDERSTDSQCSSEDEDIFFETAOVSPRGKEKQWRAMPPLSVQVSNADWW 240
Qy 241 LVKRLHKLCELCNNYIQMHLDLNCEBEPPIFKGDPFFILPFSQSSSTPSTGFSGKE 300
Db 241 LVKRLHKLCELCNNYIQMHLDLNCEBEPPIFKGDPFFILPFSQSSSTPSTGFSGKE 300
Qy 301 TPSEDRSQSREHNGESLTKAGGDDLPPSPVKEKDPKSKKEMWENGNKIYTWAD 360
Db 301 TPSEDRSQSREHNGESLTKAGGDDLPPSPVKEKDPKSKKEMWENGNKIYTWAD 360
Qy 361 KTISKMTTEYKKRQOHNLSAFPKEVVEKKGEPGPGDPSLQRPQHLMDQGQRHS 420
Db 361 KTISKMTTEYKKRQOHNLSAFPKEVVEKKGEPGPGDPSLQRPQHLMDQGQRHS 420
Qy 421 FSAGPELLRQDKRPRSGSTGSSLSVSVRDAEAOIQAWTNVLTVLNOIQILPDQFTALQ 480
Db 421 FSAGPELLRQDKRPRSGSTGSSLSVSVRDAEAOIQAWTNVLTVLNOIQILPDQFTALQ 480
Qy 481 PAVFPCTISQLTCHVTDIRVQAVREMLGRVGRVYDIIV 518
Db 481 PAVFPCTISQLTCHVTDIRVQAVREMLGRVGRVYDIIV 518

RESULT 4
ABG09728
ID ABG09728 standard; Protein; 1807 AA.
XX AC ABG09728;
XX DT 13-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #9719.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX FN W0200175067-A2.
XX PD 11-OCT-2001.

```

PF 30-MAR-2001; 2001WO-US08631.
 XX 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS73915.
 XX
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 PS
 PS Claim 20; SEQ ID No 40087; 103pp; English.
 XX
 XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 1807 AA;
 Query Match 100.0%; Score 518; DB 22; Length 1807;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RIRAMAQVFMIDTQCSPKTPNNFPHASQCLIELPEPKENGHTKKSVSFRAIVSL 60
 DB 1290 RIRAMAQVFMIDTQCSPKTPNNFPHASQCLIELPEPKENGHTKKSVSFRAIVSL 1349
 QY 61 SHQVTLQNLVYDILLEBFVKGSPGEEKTIQVPEAKIAGFLRYISQNLAVIFDILLDSYR 120
 DB 1350 SHQVTLQNLVYDILLEBFVKGSPGEEKTIQVPEAKIAGFLRYISQNLAVIFDILLDSYR 1409
 QY 121 TARFEDTSGPKCLIKKVSIGGANLVYROSMSFNIVHALVCVLTNOETITEQYVK 180
 DB 1410 TARFEDTSGPKCLIKKVSIGGANLVYROSMSFNIVHALVCVLTNOETITEQYVK 1469
 QY 181 VLFEDDERSTDSQCCSSEDEDIFETIAQVSPPRGKEKQAMRBPILLSVQVSNADWW 240
 DB 1470 VLFEDDERSTDSQCCSSEDEDIFETIAQVSPPRGKEKQAMRBPILLSVQVSNADWW 1529
 QY 241 LVKRLHLKCMELCNVYIQHLDLENCEBPPIFKGDPFILLPSFQSSSTPTSGFGSKE 300
 DB 1530 LVKRLHLKCMELCNVYIQHLDLENCEBPPIFKGDPFILLPSFQSSSTPTSGFGSKE 1589
 QY 301 TSESDRSQSRHMGESLSKAGGDLILPPSPKYEKKQPSRKKEWMAKRTYTMAD 360
 DB 1590 TSESDRSQSRHMGESLSKAGGDLILPPSPKYEKKQPSRKKEWMAKRTYTMAD 1649
 QY 361 KTISKLMTEYKKRQOQHNLSAFPEKVEKGEPLGPRQDPLLPORPHLMDQGMRS 420

DB 1650 KTISKLMTEYKKRQOQHNLSAFPEKVEKGEPLGPRQDPLLPORPHLMDQGMRS 1709
 QY 421 FSAGBELLRODKRPRSGSTGSSLSVSVDAQAQIQAMTNMVLTYLNOIQILPDQTFALQ 480
 DB 1710 FSAGBELLRODKRPRSGSTGSSLSVSVDAQAQIQAMTNMVLTYLNOIQILPDQTFALQ 1769
 QY 481 PAVFPICISQLTCHVTDIRVRAQVREMLGRVGRVYDIIV 518
 DB 1770 PAVFPICISQLTCHVTDIRVRAQVREMLGRVGRVYDIIV 1807
 RESULT 5
 ID ABB64835
 ID ABB64835 standard; Protein; 1839 AA.
 AC ABB64835;
 DT 25-FEB-2003 (first entry)
 XX
 XX Human protein SEQ ID 495.
 DE
 XX Human; expressed sequence tag; EST;
 KW haematopoietic disorder; central nervous system disease; viral infection;
 KW peripheral nervous system disease; non-healing wound; infectious disease;
 KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;
 KW fungal infection; autoimmune disorder; coagulation disorder; neutrotic;
 KW antiallergic; antiinflammatory; immunosuppressive; neuroprotective;
 KW cytostatic; haemostatic; virucide; antibacterial; fungicide;
 KW immunostimulant; cerebroprotective.
 OS
 XX Homo sapiens.
 EN WO200259260-A2.
 XX
 PD 01-AUG-2002.
 XX
 PF 16-NOV-2001; 2001WO-US42950.
 XX
 PR 17-NOV-2000; 2000US-0714936.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Goodrich RW, Liu C, Zhou P, Aundi V, Zhang J, Zhao QA;
 PI Ren F, Xue AJ, Yang Y, Weinman T, Drmanac RT;
 DR WPI; 2002-590824/63.
 DR N-PSDB; ABQ99421.
 XX
 PT New isolated polynucleotide, useful in research, diagnostic or
 PT therapeutic methods, e.g. preventing or treating disorders involving
 PT aberrant protein expression or biological activity -
 PS
 PS Claim 20; SEQ ID 495; 394pp; English.
 CC
 CC The present invention relates to novel human coding sequences
 CC (ABQ99421-ABQ99468) and proteins (ABB64832-ABB65022). The sequences are
 CC useful in therapeutic, diagnostic and research methods. The
 CC polynucleotides may be used in the field of molecular biology as
 CC hybridisation probes, primers for PCR, for chromosome and gene mapping,
 CC for the recombinant production of protein, or in generation of anti-sense
 CC DNA or RNA. The polynucleotides are useful in diagnostics as expressed
 CC sequence tags (ESTs) for identifying expressed genes or for physical
 CC mapping of the human genome. The proteins may be used as molecular weight
 CC markers, or as nutritional sources or supplements. The proteins may be
 CC used to maintain and expand cell population in a totipotent or
 CC pluripotent state useful for re-engineering damaged or diseased
 CC tissues, transplantation, manufacture of bio-pharmaceuticals or the
 CC development of bio-sensors. The polynucleotides and proteins are useful
 CC for preventing, treating or ameliorating disorders involving aberrant
 CC protein expression or biological activity, e.g. haematopoietic disorders,
 CC central/peripheral nervous system diseases, mechanical and traumatic
 CC disorders, non-healing wounds, immune deficiencies and disorders,
 CC infectious diseases caused by viral, bacterial or fungal infection.

CC autoimmune disorders, allergic reactions and conditions, coagulation
 CC disorder, or cancer. The polynucleotide sequences of the invention were
 CC assembled from ESTs isolated mainly by sequencing by hybridisation, and
 CC in some cases, sequences obtained from one or more public databases.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1839 AA;

Query Match 100.0%; Score 518; DB 23; Length 1839;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIRMAQOVFMDQCSPTNNFDHQSCLLIEPDEKNGHTKKSVSFRIVSL 60
 DB 1322 RIRMAQOVFMDQCSPTNNFDHQSCLLIEPDEKNGHTKKSVSFRIVSL 1361
 QY 61 SHQVLLQNLVDILLEEFVKGPSPEEKTIQVPEAKLAGFLRYISMQLAVIFDILLDSYR 120
 DB 1382 SHQVLLQNLVDILLEEFVKGPSPEEKTIQVPEAKLAGFLRYISMQLAVIFDILLDSYR 1441
 QY 121 TARFDTSPGLKLLKKVSGIGGANLYRQSMSPNIFHALVCAVLTNOETITAEQYKK 180
 DB 1442 TARFDTSPGLKLLKKVSGIGGANLYRQSMSPNIFHALVCAVLTNOETITAEQYKK 1501
 QY 181 VLFEDDRSTDSQOCSEDEDIPEETAQVSPRGKEKORAMPLISVQPSNADVMW 240
 DB 1502 VLFEDDRSTDSQOCSEDEDIPEETAQVSPRGKEKORAMPLISVQPSNADVMW 1561
 QY 241 LVKRLKLCMELCNNTYIQMHLLENCEBEPPIFKGDPPFILPSFQSESTPTSGFSGKE 300
 DB 1562 LVKRLKLCMELCNNTYIQMHLLENCEBEPPIFKGDPPFILPSFQSESTPTSGFSGKE 1621
 QY 301 TPSSDDDSQSRHNGESLSLKAAGGDLPLPSPKVEKDPBRKKEMWENAGKITYMAAD 360
 DB 1622 TPSSDDDSQSRHNGESLSLKAAGGDLPLPSPKVEKDPBRKKEMWENAGKITYMAAD 1681
 QY 361 KTISKLTETKKRQOHNLSAFPKVEKKEGEPPLPGPGQDSPLLQRPQHLMDQGMKHS 420
 DB 1682 KTISKLTETKKRQOHNLSAFPKVEKKEGEPPLPGPGQDSPLLQRPQHLMDQGMKHS 1741
 QY 421 FSAPPELLRODKRPRSGSTGSSLSVSVDAEAQIQAWTMVLTVLNQIQLPDQFTALQ 480
 DB 1742 FSAPPELLRODKRPRSGSTGSSLSVSVDAEAQIQAWTMVLTVLNQIQLPDQFTALQ 1801
 QY 481 PAVPFCISQLTCHVTDIRVROAVREMLGRVGRVYDIIV 518
 DB 1802 PAVPFCISQLTCHVTDIRVROAVREMLGRVGRVYDIIV 1839

RESULT 6

ABG09731

ID ABG09731 standard; Protein; 1982 AA.

AC ABG09731;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #9722.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN MO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

PI Drammac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

XX N-PSDB; AAS73918.

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 PS Claim 20; SEQ ID No 40090; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1982 AA;

Query Match 87.1%; Score 451; DB 22; Length 1982;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 451; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 REIVSLSHQVLLQNLVDILLEEFVKGPSPEEKTIQVPEAKLAGFLRYISMQLAVIF 112
 DB 1456 REIVSLSHQVLLQNLVDILLEEFVKGPSPEEKTIQVPEAKLAGFLRYISMQLAVIF 1515
 QY 113 DLLLDSYRTAREPDTSPGLKLLKKVSGIGGANLYRQSMSPNIFHALVCAVLTNOET 172
 DB 1516 DLLLDSYRTAREPDTSPGLKLLKKVSGIGGANLYRQSMSPNIFHALVCAVLTNOET 1575
 QY 173 ITAEQYKKVLFEDDERSTDSQOCSEDEDIPEETAQVSPRGKEKORAMPLISVQ 232
 DB 1576 ITAEQYKKVLFEDDERSTDSQOCSEDEDIPEETAQVSPRGKEKORAMPLISVQ 1635
 QY 233 VSNADWVWLKRLKLCMELCNNTYIQMHLLENCEBEPPIFKGDPPFILPSFQSESTPT 292
 DB 1636 VSNADWVWLKRLKLCMELCNNTYIQMHLLENCEBEPPIFKGDPPFILPSFQSESTPT 1695
 QY 293 TGGFSGKETPSEDDRSQSRHNGESLSLKAAGGDLPLPSPKVEKDPBRKKEMWENAGN 352
 DB 1696 TGGFSGKETPSEDDRSQSRHNGESLSLKAAGGDLPLPSPKVEKDPBRKKEMWENAGN 1755
 QY 353 KIYTMADKTIISKLTETKKRQOHNLSAFPKVEKKEGEPPLPGPGQDSPLLQRPQHL 412
 DB 1756 KIYTMADKTIISKLTETKKRQOHNLSAFPKVEKKEGEPPLPGPGQDSPLLQRPQHL 1815
 QY 413 DOGOMRHSFSAPPELLRODKRPRSGSTGSSLSVSVDAEAQIQAWTMVLTVLNQIQLP 472
 DB 1816 DOGOMRHSFSAPPELLRODKRPRSGSTGSSLSVSVDAEAQIQAWTMVLTVLNQIQLP 1875
 QY 473 DQFTALQPAVFPFCISQLTCHVTDIRVROAV 503

```

Db      1876 DDTFALQPAVFPICISQLTCHVTDIRVROAV 1906
|||||
RESULT 7
ID      AAY48248 standard; Protein; 192 AA.
XX      AAY48248;
AC      AAY48248;
XX      08-DEC-1999 (first entry)
DT      08-DEC-1999 (first entry)
XX      Human prostate cancer-associated protein 34.
DE      Human prostate cancer-associated protein 34.
XX      Expressed sequence tag; EST; prostate tumor; antitumor; treatment;
KW      gene therapy; tissue specificity human.
XX      Homo sapiens.
OS      Homo sapiens.
XX      DE19811193-A1.
XX      16-SEP-1999.
XX      10-MAR-1998; 98DE-1011193.
XX      10-MAR-1998; 98DE-1011193.
XX      10-MAR-1998; 98DE-1011193.
XX      (META-) METAGEN GES GENOMFORSCHUNG MBH.
XX      Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosenthal A;
XX      WPI; 1999-519628/44.
XX      N-PSDB; AA233451.
XX      New nucleic acid expressed at high level in prostatic tumor tissue and
PT      encoded polypeptides, useful for treating cancer and screening for
PT      therapeutic agents -
XX      Claim 22; 128; 166pp; German.
XX      This invention describes novel nucleic acid sequences (A) that are
CC      expressed at high level in prostatic tumor tissue and encode gene
CC      products or their fragments. The products of the invention have
CC      antitumor activity. Polypeptides (I) encoded by (A) are used: (i) for
CC      identifying agents for treatment of prostatic cancer and (ii) for
CC      therapy of prostate cancer, optionally where expressed by gene therapy
CC      methods. (A) is also used to isolate full-length genes (for gene therapy)
CC      and for recombinant production of (I), which can be used to raise
CC      specific antibodies. (A) are identified by assembly of ESTs (expressed
CC      sequence tags) before they are analyzed for expression pattern (tissue
CC      specificity). This approach eliminates many of the false results, as
CC      regards tissue specificity, associated with known methods that use
CC      single (usually short) ESTs. AAY48215-Y48303 represent protein fragments
CC      encoded by the expressed sequence tags described in the method of the
CC      invention.
XX      Sequence 192 AA;
SQ
Query Match      37.1%; Score 192; DB 20; Length 192;
Best Local Similarity 100.0%; Pred. No. 7.8e-182;
Matches 192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy      185 DDERSTDSQOC 196
|||
Db      181 DDERSTDSQOC 192
|||
RESULT 8
ID      AAY48216 standard; Protein; 180 AA.
XX      AAY48216;
AC      AAY48216;
XX      08-DEC-1999 (first entry)
DT      08-DEC-1999 (first entry)
XX      Human prostate cancer-associated protein 2.
DE      Human prostate cancer-associated protein 2.
XX      Expressed sequence tag; EST; prostate tumor; antitumor; treatment;
KW      gene therapy; tissue specificity human.
XX      Homo sapiens.
OS      Homo sapiens.
XX      DE19811193-A1.
XX      16-SEP-1999.
XX      10-MAR-1998; 98DE-1011193.
XX      10-MAR-1998; 98DE-1011193.
XX      10-MAR-1998; 98DE-1011193.
XX      (META-) METAGEN GES GENOMFORSCHUNG MBH.
XX      Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosenthal A;
XX      WPI; 1999-519628/44.
XX      N-PSDB; AA233424.
XX      New nucleic acid expressed at high level in prostatic tumor tissue and
PT      encoded polypeptides, useful for treating cancer and screening for
PT      therapeutic agents -
XX      Claim 22; 112-113; 166pp; German.
XX      This invention describes novel nucleic acid sequences (A) that are
CC      expressed at high level in prostatic tumor tissue and encode gene
CC      products or their fragments. The products of the invention have
CC      antitumor activity. Polypeptides (I) encoded by (A) are used: (i) for
CC      identifying agents for treatment of prostatic cancer and (ii) for
CC      therapy of prostate cancer, optionally where expressed by gene therapy
CC      methods. (A) is also used to isolate full-length genes (for gene therapy)
CC      and for recombinant production of (I), which can be used to raise
CC      specific antibodies. (A) are identified by assembly of ESTs (expressed
CC      sequence tags) before they are analyzed for expression pattern (tissue
CC      specificity). This approach eliminates many of the false results, as
CC      regards tissue specificity, associated with known methods that use
CC      single (usually short) ESTs. AAY48215-Y48303 represent protein fragments
CC      encoded by the expressed sequence tags described in the method of the
CC      invention.
XX      Sequence 180 AA;
SQ
Query Match      34.7%; Score 180; DB 20; Length 180;
Best Local Similarity 100.0%; Pred. No. 5.9e-170;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```


Db 121 LKCLIKKVGSGIGAGANLYRQASMSFNIFYHALVCAVLTNGETITAEQVKVLPFEDDERST 180

RESULT 9

AAO19169 standard; Protein; 49 AA.

AAO19169;

27-NOV-2002 (first entry)

Human prostate-specific PS118 protein fragment #5.

Human: prostate; prostate-specific sequence; prostate cancer; PS118;

EST; expressed sequence tag; cytostatic; gene therapy.

Homo sapiens.

US2002086316-A1.

04-JUL-2002.

26-NOV-2001; 2001US-0991681.

23-APR-1998; 98US-0065383.

23-APR-1997; 97US-0842385.

(BILL/) BILLINGEL P A.

(COHE/) COHEN M.

(COLP/) COLPITTS T L.

(FRIE/) FRIEDMAN P N.

(GORD/) GORDAN J.

(GRAN/) GRANADOS E N.

(HODG/) HODGES S C.

(KLAS/) KLAS M R.

(KRAT/) KRATOCHVIL J D.

(ROBE/) ROBERTS-RAPP L.

(RUSSE/) RUSSELL J C.

(STROU/) STROUPE S D.

Billingsel PA, Cohen M, Colpitts TL, Friedman PN, Jordan J,

Granados EN, Hodges SC, Kلاس MR, Kratochvil JD, Roberts-Rapp L;

Russell JC, Stroupe SD;

WPI; 2002-665429/71.

Novel PS118 polypeptide for detecting, diagnosing, staging, monitoring,

prognosticating, preventing, treating, or determining predisposition of

individual to diseases and conditions of prostate, e.g. prostate

cancer -

Claim 17, Page 44; 58pp; English.

The present invention relates to a number of prostate-specific sequences

derived from the human PS118 gene. These can be used in the detection,

monitoring and treatment of prostate diseases, particularly prostate

cancer. The present sequence is a PS118 protein fragment of the

invention. The coding sequences of the invention were isolated from a

prostate tissue expressed sequence tag (EST) library.

Sequence 49 AA;

Query Match 9.5%; Score 49; DB 23; Length 49;

Best Local Similarity 100.0%; Pred. No. 1.8e-40;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

393 EPLGPRGDSPLLPORPOHAMPDQGWHSFSAPELLRQDKPRSGSTGS 441

1 EPLGPRGDSPLLPORPOHAMPDQGWHSFSAPELLRQDKPRSGSTGS 49

RESULT 10

AAM50813
ID AAM50813 standard; Protein; 49 AA.

AAM50813;

01-MAY-2002 (first entry)

PS118 prostate marker immunogenic polypeptide.

PS118, prostate; marker; prostate cancer; tumour; metastasis;

benign prostatic hyperplasia; prostatic intraepithelial neoplasia;

prostaticitis; human; diagnosis; therapy; vaccine; immunogen.

Homo sapiens.

US2001055758-A1.

27-DEC-2001.

23-APR-1998; 98US-0065383.

23-APR-1997; 97US-0842385.

(BILL/) BILLING-MEDEL P A.

(COHE/) COHEN M.

(COLP/) COLPITTS T L.

(FRIE/) FRIEDMAN P N.

(GORD/) GORDAN J.

(GRAN/) GRANADOS E N.

(HODG/) HODGES S C.

(KLAS/) KLAS M R.

(KRAT/) KRATOCHVIL J D.

(ROBE/) ROBERTS-RAPP L.

Billings-Medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J,

Granados EN, Hodges SC, Kلاس MR, Kratochvil JD, Roberts-Rapp L;

Russell JC, Stroupe SD;

WPI; 2002-187683/24.

Detecting presence of target PS118 polynucleotide in test sample,

useful for detecting, diagnosing, staging, monitoring, prognosticating,

preventing or treating or determining predisposition to prostate

disease -

Claim 17, Page 44; 57pp; English.

The present sequence is that of an immunogenic polypeptide

comprising amino acids 393-441 of human prostate-specific PS118

polypeptide (see AAM50809). A PS118 consensus sequence (see

ABA91651) is found at least 12 times more often in prostate than

in non-prostate tissue. PS118 polypeptides, polynucleotides,

antibodies, agonists and inhibitors are useful for detecting,

diagnosing, staging, monitoring, prognosticating, preventing, and

treating (including by genetic immunisation), or determining the

predisposition of an individual to, diseases and conditions of the

prostate, such as benign prostatic hyperplasia, prostaticitis,

prostatic intraepithelial neoplasia, prostate cancer, tumours and

metastases. The PS118 polypeptides can be produced by expression

of PS118 polynucleotides in transfected host cells, and

immunogenic peptides are useful for raising PS118-specific

antibodies of diagnostic use. The methods and reagents of the

invention may provide an early means of detecting diseases of the

prostate and may also provide new markers which can differentiate

between the clinically important and unimportant prostate cancers

without the use of surgery.

Sequence 49 AA;

Query Match 9.5%; Score 49; DB 23; Length 49;

Best Local Similarity 100.0%; Pred. No. 1.8e-40;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 393 EPILGRGDSPLLRPOHLMDOGMHRSFASAGPELLRODKRPSGSGTGS 441
DB 1 EPLGRGDSPLLRPOHLMDOGMHRSFASAGPELLRODKRPSGSGTGS 49

RESULT 11
ID ABOG9729 standard; Protein; 50 AA.
AC ABOG9729;
DT 13-FEB-2002 (first entry)
DE Novel human diagnostic protein #9720.
KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW Food supplement; medical imaging; diagnostic; genetic disorder.
XX Homo sapiens.
OS WO200175067-A2.
PN 11-OCT-2001.
PF 30-MAR-2001; 2001WO-US08631.
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
PA (HYSE-) HYSEQ INC.
PI Drmanac RT, Liu C, Tang YT;
XX WPI: 2001-639362/73.
DR N-PSDB; AAS73916.
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX Claim 20; SEQ ID No 40088; 103bp; English.
PS The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences (I) is useful as hybridization probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABOG0010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pat_sequences.
XX
XX Sequence 50 AA;

Query Match 8.3%; Score 43; DB 22; Length 50;
Best Local Similarity 100.0%; Pred. No. 1.6e-34;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 460 MVLTVANOIQLPDQFTALQPAVFPICISQLTCHVTDIRVROA 502

4.

DB 1 MVLTVANOIQLPDQFTALQPAVFPICISQLTCHVTDIRVROA 43

RESULT 12
ID AAO19166 standard; Protein; 41 AA.
AC AAO19166;
DT 27-NOV-2002 (first entry)
DE Human prostate-specific PS118 protein fragment #2.
KM Human; prostate; prostate-specific sequence; prostate cancer; PS118;
KW EST; expressed sequence tag; cytostatic; gene therapy.
XX Homo sapiens.
OS US2002086316-A1.
PN 04-JUL-2002.
PF 26-NOV-2001; 2001US-0991681.
PR 23-APR-1998; 98US-0065383.
PR 23-APR-1997; 97US-0842385.
PA (BILL/) BILLINGEL P A.
PA (COHE/) COHEN M.
PA (COLP/) COLPITTS T L.
PA (FRIE/) FRIEDMAN P N.
PA (GORO/) GORDAN J E.
PA (GRAN/) GRANADOS E N.
PA (HODG/) HODGES S C.
PA (KLAS/) KLAAS M R.
PA (KRAT/) KRATOCHVIL J D.
PA (ROBE/) ROBERTS-RAPP L.
PA (RUSSE/) RUSSELL J C.
PA (STRO/) STROUPE S D.
XX Billengel PA, Cohen M, Colpitts TL, Friedman PN, Gordan J;
PI Granados EN, Hodges SC, Klaas MR, Kratochvil JD, Roberts-Rapp L;
PI Russell JC, Stroupe SD;
XX WPI: 2002-665429/71.
DR Novel PS118 polypeptide for detecting, diagnosing, staging, monitoring,
PT prognosticating, preventing, treating, or determining predisposition of
PT individual to diseases and conditions of prostate, e.g. prostate
PT cancer -
XX Claim 17; Page 43-44; 58bp; English.
PS The present invention relates to a number of prostate-specific sequences
XX derived from the human PS118 gene. These can be used in the detection,
XX monitoring and treatment of prostate diseases, particularly prostate
CC cancer. The present sequence is a PS118 protein fragment of the
CC invention. The coding sequences of the invention were isolated from a
CC prostate tissue expressed sequence tag (EST) library.
XX
XX Sequence 41 AA;

Query Match 7.9%; Score 41; DB 23; Length 41;
Best Local Similarity 100.0%; Pred. No. 1.3e-32;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 184 EDDERSTDSQCCSDEDDIPEETAOVSPPRGKXKORAR 224
DB 1 EDDERSTDSQCCSDEDDIPEETAOVSPPRGKXKORAR 41

RESULT 13
AAM50810

ID AAM50810 standard; Protein; 41 AA.
 AC AAM50810;
 XX
 DT 01-MAY-2002 (first entry)
 XX
 DE PS118 prostate marker immunogenic polypeptide.
 XX
 KM PS118; prostate; marker; prostate cancer; tumour; metastasis;
 KM benign prostatic hyperplasia; prostatic intraepithelial neoplasia;
 KM prostatic; human; diagnosis; therapy; vaccine; immunogen.
 XX
 OS Homo sapiens.
 XX
 PN US2001055758-A1.
 XX
 PD 27-DEC-2001.
 XX
 PF 23-APR-1998; 98US-0065383.
 XX
 PR 23-APR-1997; 97US-0842385.
 XX
 PA (BILL/) BILLING-MEDEL P A.
 PA (COHE/) COHEN M.
 PA (COLP/) COLPITTS T L.
 PA (FRIE/) FRIEDMAN P N.
 PA (GORD/) GORDON J.
 PA (GRAN/) GRANADOS E N.
 PA (HODG/) HODGES S C.
 PA (KLAS/) KLAAS M R.
 PA (KRAT/) KRATOCHVIL J D.
 PA (ROBE/) ROBERTS-RAPP L.
 XX
 PI Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;
 PI Granados EN, Hodges SC, Klaas MR, Kratochvil JD, Roberts-Rapp L;
 PI Russell JC, Stroupe SD;
 XX
 DR WPI; 2002-187683/24.
 XX
 PT Detecting presence of target PS118 polynucleotide in test sample,
 PT useful for detecting, diagnosing, staging, monitoring, prognosticating,
 PT preventing or treating or determining predisposition to prostate
 PT disease -
 XX
 PS Claim 17; Page 43; 57pp; English.
 XX
 CC The present sequence is that of an immunogenic polypeptide
 CC comprising amino acids 184-224 of human prostate-specific PS118
 CC polypeptide (see AAM50809). A PS118 consensus sequence (see
 CC ABA91651) is found at least 12 times more often in prostate than
 CC in non-prostate tissue. PS118 polypeptides, polynucleotides,
 CC antibodies, agonists and inhibitors are useful for detecting,
 CC diagnosing, staging, monitoring, prognosticating, preventing and
 CC treating (including by genetic immunisation), or determining the
 CC predisposition of an individual to, diseases and conditions of the
 CC prostate, such as benign prostatic hyperplasia, prostaticitis,
 CC prostatic intraepithelial neoplasia, prostate cancer, tumours and
 CC metastases. The PS118 polypeptides can be produced by expression
 CC of PS118 polynucleotides in transfected host cells, and
 CC immunogenic peptides are useful for raising PS118-specific
 CC antibodies of diagnostic use. The methods and reagents of the
 CC invention may provide an early means of detecting diseases of the
 CC prostate and may also provide new markers which can differentiate
 CC between the clinically important and unimportant prostate cancers
 CC without the use of surgery.
 CC
 XX
 SQ Sequence 41 AA;
 XX
 QY Query Match 7.9%; Score 41; DB 23; Length 41;
 Best Local Similarity 100.0%; Pred. No. 1.3e-23;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 EDDERSTDSQCCSSEDEDFEETAQVSPRGKRWRRAR 41
 |||
 RESULT 14
 ID AAO19168 standard; Protein; 40 AA.
 XX
 AC AAO19168;
 XX
 DT 27-NOV-2002 (first entry)
 XX
 DE Human prostate-specific PS118 protein fragment #4.
 XX
 KM Human; prostate; prostate-specific sequence; prostate cancer; PS118;
 KM EST; expressed sequence tag; cytosolic; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN US2002086316-A1.
 XX
 PD 04-JUL-2002.
 XX
 PF 26-NOV-2001; 2001US-0991681.
 XX
 PR 23-APR-1998; 98US-0065383.
 PR 23-APR-1997; 97US-0842385.
 XX
 PA (BILL/) BILLINGEL P A.
 PA (COHE/) COHEN M.
 PA (COLP/) COLPITTS T L.
 PA (FRIE/) FRIEDMAN P N.
 PA (GORD/) GORDAN J.
 PA (GRAN/) GRANADOS E N.
 PA (HODG/) HODGES S C.
 PA (KLAS/) KLAAS M R.
 PA (KRAT/) KRATOCHVIL J D.
 PA (ROBE/) ROBERTS-RAPP L.
 PA (RUSS/) RUSSELL J C.
 PA (STRO/) STROUPE S D.
 XX
 PI Billingel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;
 PI Granados EN, Hodges SC, Klaas MR, Kratochvil JD, Roberts-Rapp L;
 PI Russell JC, Stroupe SD;
 XX
 DR WPI; 2002-665429/71.
 XX
 CC Novel PS118 polypeptide for detecting, diagnosing, staging, monitoring,
 CC prognosticating, preventing, treating, or determining predisposition of
 CC individual to diseases and conditions of prostate, e.g. prostate
 CC cancer -
 XX
 PS Claim 17; Page 44; 58pp; English.
 XX
 CC The present invention relates to a number of prostate-specific sequences
 CC derived from the human PS118 gene. These can be used in the detection,
 CC monitoring and treatment of prostate diseases, particularly prostate
 CC cancer. The present sequence is a PS118 protein fragment of the
 CC prostate tissue expressed sequence tag (EST) library.
 CC
 XX
 SQ Sequence 40 AA;
 XX
 QY Query Match 7.7%; Score 40; DB 23; Length 40;
 Best Local Similarity 100.0%; Pred. No. 1.2e-31;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 332 SPVKEKDPSPKKEWENAGNKIYTMADKTIKLTMEYK 371
 |||
 DB 1 SPVKEKDPSPKKEWENAGNKIYTMADKTIKLTMEYK 40

RESULT 15

AAM50812
 ID AAM50812 standard; Protein; 40 AA.
 XX
 AC AAM50812;
 XX
 DT 01-MAY-2002 (first entry)
 XX
 DE PS118 prostate marker immunogenic polypeptide.
 XX
 KW PS118; prostate; marker; prostate cancer; tumour; metastasis;
 KW benign prostatic hyperplasia; prostatic intraepithelial neoplasia;
 KW prostatitis; human; diagnosis; therapy; vaccine; immunogen.
 XX
 OS Homo sapiens.
 XX
 PN US200105758-A1.
 XX
 PD 27-DEC-2001.
 XX
 PF 23-APR-1998; 98US-0065383.
 XX
 PR 23-APR-1997; 97US-0842385.
 XX
 PA (BILL/) BILLING-MEDEL P A.
 PA (COHE/) COHEN M.
 PA (COPL/) COPLITTIS T L.
 PA (FRIE/) FRIEDMAN P N.
 PA (GORD/) GORDON J.
 PA (GRAN/) GRANADOS E N.
 PA (HODG/) HODGES S C.
 PA (KLAS/) KLAS M R.
 PA (KRAT/) KRATOCHVIL J D.
 PA (ROBE/) ROBERTS-RAPP L.
 XX
 PI Billing-Medel PA, Cohen M, Coplittis TL, Friedman PN, Gordon J;
 PI Granados EN, Hodges SC, KLAS MR, Kratochvil JD, Roberts-Rapp L;
 PI Russell JC, Stroupe SD;
 XX
 DR WPI: 2002-187683/24.
 XX
 PT Detecting presence of target PS118 polynucleotide in test sample,
 PT useful for detecting, diagnosing, staging, monitoring, prognosticating,
 PT preventing or treating or determining predisposition to prostate
 PT disease -
 XX
 PS Claim 17; Page 44; 57pp; English.
 XX
 CC The present sequence is that of an immunogenic polypeptide
 CC comprising amino acids 332-371 of human prostate-specific PS118
 CC polypeptide (see AAM50809). A PS118 consensus sequence (see
 CC ABA91651) is found at least 12 times more often in prostate than
 CC in non-prostate tissue. PS118 polypeptides, polynucleotides,
 CC antibodies, agonists and inhibitors are useful for detecting,
 CC diagnosing, staging, monitoring, prognosticating, preventing and
 CC treating (including by genetic immunisation), or determining the
 CC predisposition of an individual to, diseases and conditions of the
 CC prostate, such as benign prostatic hyperplasia, prostatitis,
 CC prostatic intraepithelial neoplasia, prostate cancer, tumours and
 CC metastases. The PS118 polypeptides can be produced by expression
 CC of PS118 polynucleotides in transfected host cells, and
 CC immunogenic peptides are useful for raising PS118-specific
 CC antibodies of diagnostic use. The methods and reagents of the
 CC invention may provide an early means of detecting diseases of the
 CC prostate and may also provide new markers which can differentiate
 CC between the clinically important and unimportant prostate cancers
 CC without the use of surgery.
 CC
 SQ Sequence 40 AA;
 Query Match 7.7%; Score 40; DB 23; Length 40;
 Best Local Similarity 100.0%; Pred. No. 1.2e-31;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 332 SPKVEKKDPSRKKEWENAGNKIYTMAADKTIISKLMTEYK 371
 Db 1 SPKVEKKDPSRKKEWENAGNKIYTMAADKTIISKLMTEYK 40
 Search completed: July 24, 2003, 12:21:13
 Job time : 133.965 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using SW model

Run on: July 24, 2003, 12:18:29 ; Search time 59.9151 Seconds
(without alignments)
365.801 Million cell updates/sec

Title: US-09-991-681-27

Perfect score: 518
Sequence: 1 RIRAMQGVEMLDTCSPKT.....VROAVREMLGRVGRVYDIIV 518

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310858 residues

Word size : 0

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
- 5: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep:*
- 6: /cgn2_6/ptodata/1/iaa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	518	100.0	518	4	US-09-065-383-27
2	49	9.5	49	4	US-09-065-383-31
3	41	7.9	41	4	US-09-065-383-28
4	40	7.7	40	4	US-09-065-383-30
5	35	6.8	35	4	US-09-065-383-29
6	7	1.4	55	4	US-09-314-268-147
7	7	1.4	107	1	US-07-634-278-34
8	7	1.4	107	1	US-08-477-728-34
9	7	1.4	107	1	US-08-474-040-34
10	7	1.4	107	1	US-08-487-200-34
11	7	1.4	107	3	US-08-484-537-34
12	7	1.4	108	3	US-08-752-693A-2
13	7	1.4	113	1	US-08-442-542-12
14	7	1.4	113	3	US-08-765-469-12
15	7	1.4	127	3	US-08-284-516C-37
16	7	1.4	127	3	US-08-284-516C-41
17	7	1.4	127	3	US-08-284-516C-47
18	7	1.4	127	3	US-08-284-516C-57
19	7	1.4	127	3	US-08-284-516C-65
20	7	1.4	127	3	US-09-537-911A-37
21	7	1.4	127	4	US-09-537-911A-41
22	7	1.4	127	4	US-09-537-911A-47
23	7	1.4	127	4	US-09-537-911A-57
24	7	1.4	127	4	US-09-537-911A-65
25	7	1.4	128	2	US-08-449-287-6
26	7	1.4	128	2	US-08-449-287-8
27	7	1.4	138	4	US-08-976-183A-46

28	7	1.4	171	4	US-09-252-991A-29708	Sequence 29708, A
29	7	1.4	183	4	US-09-252-991A-17084	Sequence 17084, A
30	7	1.4	182	4	US-09-252-991A-31422	Sequence 31422, A
31	7	1.4	206	6	5189147-9	Patent No. 5189147
32	7	1.4	225	4	US-09-328-352-4310	Sequence 27, Appl
33	7	1.4	228	4	US-08-944-483-28	Sequence 4310, Ap
34	7	1.4	251	3	US-09-252-991A-26679	Sequence 28, Appl
35	7	1.4	282	4	US-09-107-532A-6376	Sequence 2679, A
36	7	1.4	285	4	US-09-252-991A-24032	Sequence 6376, Ap
37	7	1.4	301	4	US-09-198-452A-205	Sequence 24032, A
38	7	1.4	313	4	US-09-252-991A-31554	Sequence 205, App
39	7	1.4	325	4	US-09-252-991A-23334	Sequence 31554, A
40	7	1.4	367	4	US-09-220-528-63	Sequence 23334, A
41	7	1.4	400	3	US-09-252-991A-31461	Sequence 63, Appl
42	7	1.4	405	4	US-09-308-345A-4	Sequence 31461, A
43	7	1.4	513	4	US-09-308-345A-2	Sequence 4, Appl1
44	7	1.4	640	4	US-09-308-345A-2	Sequence 2, Appl1
45	7	1.4	659	4	US-09-252-991A-19154	Sequence 19154, A

ALIGNMENTS

RESULT 1
US-09-065-383-27
Sequence 27, Application US/09065383
Patent No. 6391543
GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GORDON, JULIAN
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KLAAS, MICHAEL R.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: ROBERTS-RAP, LISA
APPLICANT: ROBERTS, JOHN C.
APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
TITLE OF INVENTION: FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/065,383
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/842,385
FILING DATE: 23-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6084, US, P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 518 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6391543e
US-09-065-383-27

Query Match 100.0%; Score 518; DB 4; Length 518;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIRAMAQVFMIDTCCSPKTPNNPFAOCCOLIIEIPREKNGHTKTSVSPREIIVSL 60
DB 1 RIRAMAQVFMIDTCCSPKTPNNPFAOCCOLIIEIPREKNGHTKTSVSPREIIVSL 60
QY 61 SHOVLQNLVYDILLEEYVKGPSGEEKTIQVEAKLAGLRYSIMONTAVIPDLLDSYR 120
DB 61 SHOVLQNLVYDILLEEYVKGPSGEEKTIQVEAKLAGLRYSIMONTAVIPDLLDSYR 120
QY 121 TAREEDTSPGLKCLIKYVSGIGGANIVKQASMSFNIVYHALVCAVLNQETITAEQYK 180
DB 121 TAREEDTSPGLKCLIKYVSGIGGANIVKQASMSFNIVYHALVCAVLNQETITAEQYK 180
QY 181 VLFEDDERSTDSQCCSSEDEDIPEETAQVSPRGKEKQKQARAPPLSVQVSNADWV 240
DB 181 VLFEDDERSTDSQCCSSEDEDIPEETAQVSPRGKEKQKQARAPPLSVQVSNADWV 240
QY 241 LVKRLHLKLMELCNVYIQMHLDLENCMEBPPIFKGDPFILLPSQSESTPSTGFSGKE 300
DB 241 LVKRLHLKLMELCNVYIQMHLDLENCMEBPPIFKGDPFILLPSQSESTPSTGFSGKE 300
QY 301 TPSEDDRSQSRHEMESLSLKAQGGDILLPSPKYEKQKQKQKQKQKQKQKQKQKQKQ 360
DB 301 TPSEDDRSQSRHEMESLSLKAQGGDILLPSPKYEKQKQKQKQKQKQKQKQKQKQKQ 360
QY 361 KTISKLMEYKQ 420
DB 361 KTISKLMEYKQ 420
QY 421 FSAGPELLRQDKRPPSGSTGSSLSVSRDAEAQIQAMTNVLTIVNQIQLPDQFTALQ 480
DB 421 FSAGPELLRQDKRPPSGSTGSSLSVSRDAEAQIQAMTNVLTIVNQIQLPDQFTALQ 480
QY 481 PAVFPCISQLTCHVDIRVQAVREMLGRVGRVYDIIV 518
DB 481 PAVFPCISQLTCHVDIRVQAVREMLGRVGRVYDIIV 518

RESULT 2
US-09-065-383-31
Sequence 31, Application US/09065383
Patent No. 6391543

GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GORDON, JULIAN
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KLAS, MICHAEL R.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: ROBERTS-RAPP, LISA
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSES: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA

ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/065,383
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/842,385
FILING DATE: 23-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6084, US, P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6391543e
US-09-065-383-31

Query Match 9.5%; Score 49; DB 4; Length 49;
Best Local Similarity 100.0%; Pred. No. 4,1e-41;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 EPLPRGQDSPILQRPQHLMDQGMHSFSAGPELLRQDKRPPSGSTGS 441
DB 1 EPLPRGQDSPILQRPQHLMDQGMHSFSAGPELLRQDKRPPSGSTGS 49

RESULT 3
US-09-065-383-28
Sequence 28, Application US/09065383
Patent No. 6391543

GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GORDON, JULIAN
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KLAS, MICHAEL R.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: ROBERTS-RAPP, LISA
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSES: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/065,383

FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/842,385
FILING DATE: 23-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6084.US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 41 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6391543e
US-09-065-383-28

Query Match 7.9%; Score 41; DB 4; Length 41;
Best Local Similarity 100.0%; Pred. No. 3.1e-33;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 184 EDDERSTDSSQCCSEDEDEFEETAAOVSPRKGKRRAR 224
Db 1 EDDERSTDSSQCCSEDEDEFEETAAOVSPRKGKRRAR 41

RESULT 4
US-09-065-383-30
Sequence 30, Application US/09065383
Patent No. 6391543
GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GORDON, JULIAN
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KLAS, MICHAEL R.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: ROBERTS-RAPP, LISA
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
NUMBER OF INVENTION: FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/065,383
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/842,385
FILING DATE: 23-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441

REFERENCE/DOCKET NUMBER: 6084.US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6391543e
US-09-065-383-30

Query Match 7.7%; Score 40; DB 4; Length 40;
Best Local Similarity 100.0%; Pred. No. 3e-32;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 332 SPKVEKKDPSRKKEWENAGNKITYMAADXTISKLTMEYK 371
Db 1 SPKVEKKDPSRKKEWENAGNKITYMAADXTISKLTMEYK 40

RESULT 5
US-09-065-383-29
Sequence 29, Application US/09065383
Patent No. 6391543
GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GORDON, JULIAN
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KLAS, MICHAEL R.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: ROBERTS-RAPP, LISA
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
NUMBER OF INVENTION: FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/065,383
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/842,385
FILING DATE: 23-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6084.US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 35 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6391543e
US-09-065-383-29

Query Match 6.8%; Score 35; DB 4; Length 35;
Best Local Similarity 100.0%; Pred. No. 2,4e-27;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 283 SFOSESSTPTGFGSGKETPSDDRSQSRHMGES 317
Db 1 SFOSESSTPTGFGSGKETPSDDRSQSRHMGES 35

RESULT 6
US-09-314-268-147
Sequence 147, Application US/09314268
Patent No. 6346377
GENERAL INFORMATION:

APPLICANT: Dootbar, John
TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA
FILE REFERENCE: 3789/80902
CURRENT APPLICATION NUMBER: US/09/314,268
CURRENT FILING DATE: 1999-03-19
EARLIER APPLICATION NUMBER: 09/314,268
EARLIER FILING DATE: 1999-05-18
NUMBER OF SEQ ID NOS: 179
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 147
LENGTH: 55
TYPE: PRT
ORGANISM: Human papillomavirus type 23
US-09-314-268-147

Query Match 1.4%; Score 7; DB 4; Length 55;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 PSPGEEK 87
Db 19 PSPGEEK 25

RESULT 7
US-07-634-278-34
Sequence 34, Application US/07634278
Patent No. 5530101
GENERAL INFORMATION:

APPLICANT: QUEEN, Cary L.
APPLICANT: CO, Man Sung
APPLICANT: SCHNEIDER, William P.
APPLICANT: LANDOLFI, Nicholas F.
APPLICANT: COELINGH, Kathleen L.
APPLICANT: SELICK, Harold E.
TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: 379 Lytton Avenue
CITY: Palo Alto
STATE: California
COUNTRY: US
ZIP: 94301

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/634,278
FILING DATE: 19-DEC-1990

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/590,274
FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/310,252
FILING DATE: 13-FEB-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/290,975
FILING DATE: 28-DEC-1988
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M.
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 11823-002600

TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYDROTHERICAL: NO

FEATURE:
NAME/KEY: Protein
LOCATION: 1..107
OTHER INFORMATION: /note="Amino acid sequence of the
OTHER INFORMATION: light chain for humane Lay antibody."
US-07-634-278-34

Query Match 1.4%; Score 7; DB 1; Length 107;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 441 SSISSVS 447
Db 9 SSISSVS 15

RESULT 8
US-08-477-728-34
Sequence 34, Application US/08477728
Patent No. 5585089
GENERAL INFORMATION:

APPLICANT: QUEEN, Cary L.
APPLICANT: SCHNEIDER, William P.
APPLICANT: SELICK, Harold E.
TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: Palo Alto
STATE: California
COUNTRY: US
ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,728
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/634,278
FILING DATE: 19-DEC-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/590,274
FILING DATE: 28-SEP-1990

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/310,252
FILING DATE: 13-FEB-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/290,975
FILING DATE: 28-DEC-1988
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 11823-002600
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: Protein
LOCATION: 1..107
OTHER INFORMATION:
OTHER INFORMATION: /note="Amino acid sequence of the
US-08-474-040-34 light chain for humane Lay antibody."

Query Match 1.4%; Score 7; DB 1; Length 107;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 441 SLSVSV 447
Db 9 SLSVSV 15

RESULT 9
US-08-474-040-34
Sequence 34, Application US/08474040
Patent No. 5693761
GENERAL INFORMATION:
APPLICANT: QUEEN, Cary L.
APPLICANT: CO, Man Sung
APPLICANT: SCHNEIDER, William P.
APPLICANT: LANDOLFI, Nicholas F.
APPLICANT: COELINGH, Kathleen L.
APPLICANT: SELICK, Harold E.
TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: 379 Lytton Avenue
CITY: Palo Alto
STATE: California
COUNTRY: US
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,040
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/634,278
FILING DATE: 19-DEC-1990
APPLICATION NUMBER: US 07/590,274
FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/310,252

FILING DATE: 13-FEB-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/290,975
FILING DATE: 28-DEC-1988
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 11823-002600
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: Protein
LOCATION: 1..107
OTHER INFORMATION:
OTHER INFORMATION: /note="Amino acid sequence of the
US-08-474-040-34 light chain for humane Lay antibody."

Query Match 1.4%; Score 7; DB 1; Length 107;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 441 SLSVSV 447
Db 9 SLSVSV 15

RESULT 10
US-08-487-200-34
Sequence 34, Application US/08487200
Patent No. 5693762
GENERAL INFORMATION:
APPLICANT: QUEEN, Cary L.
APPLICANT: CO, Man Sung
APPLICANT: SCHNEIDER, William P.
APPLICANT: LANDOLFI, Nicholas F.
APPLICANT: COELINGH, Kathleen L.
APPLICANT: SELICK, Harold E.
TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew
STREET: 379 Lytton Avenue
CITY: Palo Alto
STATE: California
COUNTRY: US
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,200
FILING DATE: 7-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/634,278
FILING DATE: 19-DEC-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/590,274
FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/310,252
FILING DATE: 13-FEB-1989

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/290,975
FILING DATE: 28-DEC-1988
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 11823-002610
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: Protein
LOCATION: 1..107
OTHER INFORMATION: /note="Amino acid sequence of the
US-08-487-200-34 light chain for humane lay antibody."
Query Match 1.4%; Score 7; DB 1; Length 107;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 441 SLSVS SV 447
DB 9 SLSVS SV 15
RESULT 11
US-08-484-537-34
Sequence 34, Application US/08484537
Patent No. 6180370
GENERAL INFORMATION:
APPLICANT: QUEEN, Cary L.
APPLICANT: CO, Man Sung
APPLICANT: SCHNEIDER, William P.
APPLICANT: LANDOLFI, Nicholas F.
APPLICANT: COELINGH, Kathleen L.
APPLICANT: SELICK, Harold E.
TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourile and Crew
STREET: 379 Lytton Avenue
CITY: Palo Alto
STATE: California
COUNTRY: US
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,537
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/634,278
FILING DATE: 19-DEC-1990
APPLICATION NUMBER: US 07/590,274
FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/310,252
FILING DATE: 13-FEB-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/290,975

FILING DATE: 28-DEC-1988
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 11823-002600
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: Protein
LOCATION: 1..107
OTHER INFORMATION: /note="Amino acid sequence of the
US-08-484-537-34 light chain for humane lay antibody."
Query Match 1.4%; Score 7; DB 3; Length 107;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 441 SLSVS SV 447
DB 9 SLSVS SV 15
RESULT 12
US-08-752-693A-2
Sequence 2, Application US/08752693A
Patent No. 6190640
GENERAL INFORMATION:
APPLICANT: Sydney Welt
Geord Riter
Leonard Cohen
Clarence William Jr.
Elizabeth Carswell Richards
Mary John
TITLE OF INVENTION: METHOD FOR TREATING NEOPLASIA
USING HUMANIZED ANTIBODIES WHICH
BIND TO ANTIGEN A33
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fulbright & Jaworski LLP
STREET: 666 Fifth Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10103
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette
COMPUTER: IBM PAULINE STASIAK, PH.D./2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/752,693A
FILING DATE: 19-NOV. 6190640-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/07321
FILING DATE: May 21, 1996
APPLICATION NUMBER: 08/020,223
FILING DATE: February 16, 1993
APPLICATION NUMBER: 07/573,153
FILING DATE: March 18, 1991
APPLICATION NUMBER: 07/327,765
FILING DATE: March 23, 1989
APPLICATION NUMBER: 07/118,411

FILING DATE: No. 6190640ember 6, 1987
 APPLICATION NUMBER: 06/724,991
 FILING DATE: April 19, 1985
 ATTORNEY/AGENT INFORMATION:
 NAME: Mary Anne Schofield
 REGISTRATION NUMBER: 36,669
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 318-3000
 TELEFAX: (212) 318-3400
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 108
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: <Unknown>
 DESCRIPTION: SEQ ID NO: 2
 ORGANISM: Human A33
 SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 US-08-752-693A-2

Query Match 1.4%; Score 7; DB 3; Length 108;
 Best Local Similarity 100.0%; Pred. No. 43;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 441 SSLSVSV 447
 Db 9 SSLSVSV 15

RESULT 13
 US-08-442-542-12
 Sequence 12, Application US/08442542
 Patent No. 3686600
 GENERAL INFORMATION:
 APPLICANT: Carozzi, Nadine B.
 APPLICANT: Kozziel, Michael G.
 TITLE OF INVENTION: Antibodies which bind to Insect Gut
 TITLE OF INVENTION: Proteins and their Use
 NUMBER OF SEQUENCES: 49
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Ciba-Geigy Corporation
 STREET: 7 Skyline Drive
 CITY: Hawthorne
 STATE: NY
 COUNTRY: USA
 ZIP: 10532
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/442,542
 FILING DATE: 16-MAY-1995
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/267,641
 FILING DATE: 28-JUN-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Spruill, W. Murray
 REGISTRATION NUMBER: 32,943
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 919-541-8615
 TELEFAX: 919-541-8689
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 113 amino acids
 TYPE: amino acid
 TOPOLOGY: linear

MOLECULE TYPE: protein
 US-08-442-542-12

Query Match 1.4%; Score 7; DB 1; Length 113;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 441 SSLSVSV 447
 Db 9 SSLSVSV 15

RESULT 14
 US-08-765-469-12
 Sequence 12, Application US/08765469
 Patent No. 6063301
 GENERAL INFORMATION:
 APPLICANT: Carozzi, Nadine B.
 APPLICANT: Kozziel, Michael G.
 TITLE OF INVENTION: Antibodies which bind to Insect Gut
 TITLE OF INVENTION: Proteins and their Use
 NUMBER OF SEQUENCES: 49
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Ciba-Geigy Corporation
 STREET: 7 Skyline Drive
 CITY: Hawthorne
 STATE: NY
 COUNTRY: USA
 ZIP: 10532
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/765,469
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/267,641
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Spruill, W. Murray
 REGISTRATION NUMBER: 32,943
 REFERENCE/DOCKET NUMBER: CGC 1750
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 919-541-8615
 TELEFAX: 919-541-8689
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 113 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-765-469-12

Query Match 1.4%; Score 7; DB 3; Length 113;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 441 SSLSVSV 447
 Db 9 SSLSVSV 15

RESULT 15
 US-08-284-516C-37
 Sequence 37, Application US/08284516C
 Patent No. 6056957
 GENERAL INFORMATION:
 APPLICANT: Chou, Chuan-Chu
 APPLICANT: Murgolo, Nicholas
 APPLICANT: Abrams, John

APPLICANT: Jenh, Chung-Her
APPLICANT: Petro, Mary
APPLICANT: Silver, Jon
APPLICANT: Tindall, Stephen
APPLICANT: Windsor, William
APPLICANT: Zavadny, Paul
TITLE OF INVENTION: Design, Cloning and Expression of Humanized Monoclonal Antib
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schering-Plough Corporation
STREET: 2000 Galloping Hill Road
CITY: Kenilworth
STATE: New Jersey
COUNTRY: USA
ZIP: 07033-0530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 6.0.5
SOFTWARE: Microsoft Word 4.00B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/284,516C
FILING DATE: 04-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/00759
FILING DATE: 04-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/832,842
FILING DATE: 06-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Foulke, Cynthia L.
REGISTRATION NUMBER: 32,364
REFERENCE/DOCKET NUMBER: JB0233K
TELECOMMUNICATION INFORMATION:
TELEPHONE: 908 298 2987
TELEFAX: 908 298 5388
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 127 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-284-516C-37

Query Match 1.4%; Score 7; DB 3; Length 127;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 441 SLSVSV 447
|||
Db 28 SLSVSV 34

Search completed: July 24, 2003, 12:28:36
Job time : 60.9151 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OW protein - protein search, using sw model

Run on: July 24, 2003, 12:26:04 ; Search time 87.9766 Seconds
(without alignments)
639.249 Million cell updates/sec

Title: US-09-991-681-27

Perfect score: 518
Sequence: 1 RIRAAQOVFMDTQCSPKT.....VRQAVREWLGRVHYDIIV 518

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 451899 seqs, 118759770 residues

Word size : 0

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Published Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubppaa/PCR_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubppaa/CTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep:*
- 12: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubppaa/US10C_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1.5	32	9	US-09-864-761-35805	Sequence 35805, A
2	1.5	1031	9	US-09-815-242-10932	Sequence 10932, A
3	1.4	10	11	US-09-572-4048-3170	Sequence 3170, Ap
4	1.4	21	9	US-09-864-761-41759	Sequence 41759, A
5	1.4	49	11	US-09-768-2358-36	Sequence 36, App1
6	1.4	60	9	US-09-864-761-45360	Sequence 45360, A
7	1.4	115	10	US-09-764-868-815	Sequence 815, App
8	1.4	119	11	US-09-764-891-3238	Sequence 3238, App
9	1.4	121	11	US-09-974-879-143	Sequence 143, App
10	1.4	122	11	US-09-305-736-142	Sequence 142, App
11	1.4	135	9	US-09-916-790-31	Sequence 31, App1
12	1.4	186	9	US-09-729-674-30	Sequence 30, App1
13	1.4	218	11	US-09-925-179-67	Sequence 67, App1
14	1.4	253	9	US-09-815-242-11842	Sequence 11842, A
15	1.4	259	15	US-10-128-714-3274	Sequence 3274, Ap

16	7	1.4	272	15	US-10-128-714-8274	Sequence 8274, Ap
17	7	1.4	287	15	US-10-156-761-11165	Sequence 11165, A
18	7	1.4	314	9	US-09-925-299-934	Sequence 934, App
19	7	1.4	314	11	US-09-925-299-934	Sequence 934, App
20	7	1.4	314	15	US-10-106-698-4725	Sequence 4725, Ap
21	7	1.4	320	10	US-09-738-626-6258	Sequence 6258, Ap
22	7	1.4	342	10	US-09-738-626-4170	Sequence 4170, Ap
23	7	1.4	352	15	US-10-156-761-8669	Sequence 8669, Ap
24	7	1.4	359	9	US-09-764-898-210	Sequence 210, App
25	7	1.4	365	10	US-09-895-913A-264	Sequence 264, App
26	7	1.4	400	9	US-09-920-920-63	Sequence 63, App1
27	7	1.4	400	9	US-09-828-366-16	Sequence 16, App1
28	7	1.4	405	15	US-10-156-761-12073	Sequence 12073, A
29	7	1.4	414	15	US-10-156-761-13296	Sequence 13296, A
30	7	1.4	438	10	US-09-866-582-28	Sequence 28, App1
31	7	1.4	445	9	US-09-815-242-11701	Sequence 11701, A
32	7	1.4	454	10	US-09-866-582-8	Sequence 8, App1
33	7	1.4	477	15	US-10-161-968-14	Sequence 14, App1
34	7	1.4	569	10	US-09-738-626-6179	Sequence 6179, App
35	7	1.4	578	10	US-09-999-248-2	Sequence 2, App1
36	7	1.4	581	9	US-09-925-302-801	Sequence 801, App
37	7	1.4	586	11	US-09-946-374-243	Sequence 243, App
38	7	1.4	596	14	US-10-006-867-100	Sequence 100, App
39	7	1.4	596	14	US-10-052-586-310	Sequence 310, App
40	7	1.4	596	14	US-10-063-547-100	Sequence 100, App
41	7	1.4	596	15	US-10-174-590-310	Sequence 310, App
42	7	1.4	596	15	US-10-176-758-310	Sequence 310, App
43	7	1.4	596	15	US-10-175-737-310	Sequence 310, App
44	7	1.4	596	15	US-10-063-616-100	Sequence 100, App
45	7	1.4	596	15	US-10-173-706-310	Sequence 310, App

ALIGNMENTS

RESULT 1
US-09-864-761-35805
Sequence 35805, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OR INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aeomica-X-1
CURRENT FILING DATE: 2001-05-23
CURRENT APPLICATION NUMBER: US/09/864, 761
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/180, 312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207, 456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632, 366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263, 6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236, 359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30

```

; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 35805
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007688.15
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 35
; OTHER INFORMATION: EXPRESSED IN HEAT, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 31
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.9
; OTHER INFORMATION: EST_HUMAN HIT: BE531168.1, EVALU6 3.90e-01
; US-09-864-761-35805

```

```

Query Match      1.5%; Score 8; DB 9; Length 32;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 419 HSFSAPE 426
 13 HSFSAPE 20

```

RESULT 2
; US-09-815-242-10932
; Sequence 10932, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308

```

```

; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10932
; LENGTH: 1031
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
; US-09-815-242-10932

```

```

Query Match      1.5%; Score 8; DB 9; Length 1031;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 286 SESSTPST 293
 796 SESSTPST 803

```

RESULT 3
; US-09-572-404B-3170
; Sequence 3170, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProPatent version 1.0
; SEQ ID NO 3170
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in unknown at 6-15 and may interact with sequence
; OTHER INFORMATION: in this patent.
; US-09-572-404B-3170

```

```

Query Match      1.4%; Score 7; DB 11; Length 10;
Best Local Similarity 100.0%; Pred. No. 9.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 441 SLSVSV 447
 4 SLSVSV 10

```

RESULT 4
; US-09-864-761-41759
; Sequence 41759, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecmica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666

```

```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 41759
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AF233390.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.4
; US-09-864-761-41759

Query Match      1.4%; Score 7; DB 9; Length 21;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      436 SGGTSS 442
      |||||
Db      8 SGGTSS 14

RESULT 5
US-09-768-235B-36
; Sequence 36, Application US/09768235B
; Publication No. US20030024003A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Marcus
; APPLICANT: Reindl, Andreas
; APPLICANT: Schmidt, Ralf-Michael
; APPLICANT: Freund, Annette
; APPLICANT: Ehrhardt, Thomas
; APPLICANT: Biechoff, Friedrich
; APPLICANT: Renz, Andreas
; APPLICANT: Duwens, Elke
; APPLICANT: Clippus, Petra
; APPLICANT: Lersch, Jens
; APPLICANT: Reski, Ralf
; TITLE OF INVENTION: Moss genes from Physcomitrella patens encoding proteins
; TITLE OF INVENTION: Involved in the regulation of cell division, growth and
; FILE REFERENCE: 0093/00009
; CURRENT APPLICATION NUMBER: US/09/768,235B
```

```

; CURRENT FILING DATE: 2002-03-11
; PRIOR APPLICATION NUMBER: PCT/EP00/00675
; PRIOR FILING DATE: 2000-01-28
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 2.1/WordPerfect version 6.1
; SEQ ID NO 36
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Physcomitrella patens
; US-09-768-235B-36

Query Match      1.4%; Score 7; DB 11; Length 49;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      178 VKKVLFE 184
      |||||
Db      39 VKKVLFE 45

RESULT 6
US-09-864-761-45360
; Sequence 45360, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 45360
```

LENGTH: 60
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC013626.3
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
OTHER INFORMATION: EST HUMAN HIT: F06450.1, EVALUATE 4.00e-26
OTHER INFORMATION: SWISSPROT HIT: P36607, EVALUATE 1.00e-08
US-09-864-761-45360

Query Match 1.4%; Score 7; DB 9; Length 60;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 318 LSLKAG 324
Db 38 LSLKAG 44

RESULT 7
US-09-764-868-815
Sequence 815, Application US/09764868
Patent No. US2020168711A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PT232
CURRENT APPLICATION NUMBER: US/09/764,868
CURRENT FILING DATE: 2001-01-17
Prior application data removed - refer to PALM or file wrapper
NUMBER OF SEQ ID NOS: 1510
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 815
LENGTH: 115
TYPE: PRT
ORGANISM: Homo sapiens
US-09-764-868-815

Query Match 1.4%; Score 7; DB 10; Length 115;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 425 PELLROD 431
Db 77 PELLROD 83

RESULT 8
US-09-764-891-3238
Sequence 3238, Application US/09764891
Publication No. US20030077808A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC006
CURRENT APPLICATION NUMBER: US/09/764,891
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 10231
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3238
LENGTH: 119
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (10)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

NAME/KEY: SITE
LOCATION: (19)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (125)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (45)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-891-3238

Query Match 1.4%; Score 7; DB 11; Length 119;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 327 LLLPSP 333
Db 66 LLLPSP 72

RESULT 9
US-09-974-879-143
Sequence 143, Application US/0974879
Publication No. US200302803A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 125 Human Secreted Proteins
FILE REFERENCE: P2020P2
CURRENT APPLICATION NUMBER: US/09/974,879
CURRENT FILING DATE: 2001-10-12
Prior application data removed - refer to PALM or file wrapper
Prior application number: US 60/239,893
Prior filing date: 2000-10-13
Prior application number: US 09/818,683
Prior filing date: 2001-03-28
Prior application number: US 09/305,736
Prior filing date: 1999-05-05
Prior application number: PCT/US98/23435
Prior filing date: 1998-11-04
Prior application number: US 60/064,911
Prior filing date: 1997-11-07
Prior application number: US 60/064,912
Prior filing date: 1997-11-07
Prior application number: US 60/064,983
Prior filing date: 1997-11-07
Prior application number: US 60/064,900
Prior filing date: 1997-11-07
Prior application number: US 60/064,988
Prior filing date: 1997-11-07
Prior application number: US 60/064,987
Prior filing date: 1997-11-07
Prior application number: US 60/064,908
Prior filing date: 1997-11-07
Prior application number: US 60/064,984
Prior filing date: 1997-11-07
Prior application number: US 60/064,985
Prior filing date: 1997-11-07
Prior application number: US 60/066,094
Prior filing date: 1997-11-17
Prior application number: US 60/066,100
Prior filing date: 1997-11-17
Prior application number: US 60/066,089
Prior filing date: 1997-11-17
Prior application number: US 60/066,095
Prior filing date: 1997-11-17
Prior application number: US 60/066,090
Prior filing date: 1997-11-17
Number of seq id nos: 611
Software: PatentIn Ver. 2.0
Seq id no 143
Length: 121
Type: PRT
Organism: Homo sapiens
US-09-974-879-143

Query Match 1.4%; Score 7; DB 11; Length 121;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 297 SGETPS 303
Db 70 SGETPS 76

RESULT 10
US-09-305-736-142
; Sequence 142, Application US/09305736
; Publication No. US20030088078A1
; GENERAL INFORMATION:
; APPLICANT: Peng et al.
; TITLE OF INVENTION: 125 Human Secreted Proteins
; FILE REFERENCE: P2020P1
; CURRENT FILING DATE: 1999-05-05
; EARLIER APPLICATION NUMBER: PCT/US98/23435
; EARLIER FILING DATE: 1998-11-04
; EARLIER APPLICATION NUMBER: 60/064,911
; EARLIER FILING DATE: 1997-11-07
; EARLIER APPLICATION NUMBER: 60/064,912
; EARLIER FILING DATE: 1997-11-07
; EARLIER APPLICATION NUMBER: 60/064,983
; EARLIER FILING DATE: 1997-11-07
; EARLIER APPLICATION NUMBER: 60/064,900
; EARLIER FILING DATE: 1997-11-07
; EARLIER APPLICATION NUMBER: 60/064,988
; EARLIER FILING DATE: 1997-11-07
; EARLIER APPLICATION NUMBER: 60/064,987
; EARLIER FILING DATE: 1997-11-07
; EARLIER APPLICATION NUMBER: 60/064,908
; EARLIER FILING DATE: 1997-11-07
; EARLIER APPLICATION NUMBER: 60/064,984
; EARLIER FILING DATE: 1997-11-07
; EARLIER APPLICATION NUMBER: 60/064,985
; EARLIER FILING DATE: 1997-11-07
; EARLIER APPLICATION NUMBER: 60/066,094
; EARLIER FILING DATE: 1997-11-17
; EARLIER APPLICATION NUMBER: 60/066,100
; EARLIER FILING DATE: 1997-11-17
; EARLIER APPLICATION NUMBER: 60/066,089
; EARLIER FILING DATE: 1997-11-17
; EARLIER APPLICATION NUMBER: 60/066,095
; EARLIER FILING DATE: 1997-11-17
; EARLIER APPLICATION NUMBER: 60/066,090
; EARLIER FILING DATE: 1997-11-17
; NUMBER OF SEQ ID NOS: 612
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 142
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (122)
; OTHER INFORMATION: Xaa equals stop translation
US-09-305-736-142

Query Match 1.4%; Score 7; DB 11; Length 122;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 297 SGETPS 303
Db 70 SGETPS 76

RESULT 11
US-09-916-790-31

; Sequence 31, Application US/09916790
; Patent No. US20020061573A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; APPLICANT: Kapeller-Liebermann, Rosana
; APPLICANT: Silos-Santiago, Immaculada
; TITLE OF INVENTION: 18431 AND 32374, NOVEL HUMAN PROTEIN
; TITLE OF INVENTION: KINASE FAMILY MEMBERS AND USES THEREFOR
; FILE REFERENCE: 381552002700
; CURRENT APPLICATION NUMBER: US/09/916,790
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 60/221,543
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PasteSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus amino acid sequence
US-09-916-790-31

Query Match 1.4%; Score 7; DB 9; Length 135;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 425 PELLROD 431
Db 93 PELLROD 99

RESULT 12
US-09-729-674-30
; Sequence 30, Application US/09729674
; Patent No. US2001003935A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavallee, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Agostino, Michael J.
; APPLICANT: Steininger II, Robert J.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fectel, Kim
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: 6055-64X
; CURRENT APPLICATION NUMBER: US/09/729,674
; CURRENT FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: 09/539,330
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 283
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 186
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-729-674-30

Query Match 1.4%; Score 7; DB 9; Length 186;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 286 SESSTPS 292
Db 123 SESSTPS 129

RESULT 13
US-09-925-179-67
; Sequence 67, Application US/09925179
; Publication No. US20030044858A1
; GENERAL INFORMATION:
; APPLICANT: Jardieu, Paula M.
; TITLE OF INVENTION: Anti-19E Antibodies (as amended)
; FILE REFERENCE: P0718P2CID1C1US
; CURRENT APPLICATION NUMBER: US/09/925,179
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 08/466,163
; PRIOR FILING DATE: 1995-06-06
; PRIOR APPLICATION NUMBER: US 08/405,617
; PRIOR FILING DATE: 1995-03-15
; PRIOR APPLICATION NUMBER: US 08/185,899
; PRIOR FILING DATE: 1994-01-26
; PRIOR APPLICATION NUMBER: PCT/US92/06860
; PRIOR FILING DATE: 1992-08-14
; PRIOR APPLICATION NUMBER: US 07/879,495
; PRIOR FILING DATE: 1992-05-07
; PRIOR APPLICATION NUMBER: US 07/744,768
; PRIOR FILING DATE: 1991-08-14
; NUMBER OF SEQ ID NOS: 68
; SEQ ID NO 67
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Full-length light chain sequence corresponding to F(ab)9 of Table
US-09-925-179-67

Query Match 1.4%; Score 7; DB 11; Length 218;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 441 SLSLSV 447
Db 9 SLSLSV 15

RESULT 14
US-09-815-242-11842
; Sequence 11842, Application US/09815242
; Patent No. US2002061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11842
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-11842

Query Match 1.4%; Score 7; DB 9; Length 253;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 507 LGRVGRV 513
Db 215 LGRVGRV 221

RESULT 15
US-10-128-714-3274
; Sequence 3274, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Weng
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; TITLE OF INVENTION: Methods of Use
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; PRIOR FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3274
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-10-128-714-3274

Query Match 1.4%; Score 7; DB 15; Length 259;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 326 DLPLPPS 332
Db 61 DLPLPPS 67

Search completed: July 24, 2003, 12:48:25
Job time : 91.9766 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 12:21:29 ; Search time 780.413 Seconds
(without alignments)
577.701 Million cell updates/sec

Title: US-09-991-681-27

Perfect score: 518

Sequence: 1 RIRAMAQQVEMLDTCSPKT.....VROAVHEWLGKGVYDIIV 518

Scoring table: OLIGO

Searched: 5580241 seqs, 870357830 residues

Word size: 0

Total number of hits satisfying chosen parameters: 5580241

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: Pending Patents AA Main:*

1: /cgn2_6/prodata/1/paa/US06_COMB.pep.*
2: /cgn2_6/prodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/prodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/prodata/1/paa/US08_COMB.pep.*
5: /cgn2_6/prodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/prodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/prodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/prodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/prodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/prodata/1/paa/US086_COMB.pep.*
11: /cgn2_6/prodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/prodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/prodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/prodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/prodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/prodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/prodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/prodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/prodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/prodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/prodata/1/paa/US097A_COMB.pep.*
22: /cgn2_6/prodata/1/paa/US097B_COMB.pep.*
23: /cgn2_6/prodata/1/paa/US098_COMB.pep.*
24: /cgn2_6/prodata/1/paa/US099A_COMB.pep.*
25: /cgn2_6/prodata/1/paa/US099B_COMB.pep.*
26: /cgn2_6/prodata/1/paa/US100_COMB.pep.*
27: /cgn2_6/prodata/1/paa/US101_COMB.pep.*
28: /cgn2_6/prodata/1/paa/US102_COMB.pep.*
29: /cgn2_6/prodata/1/paa/US103_COMB.pep.*
30: /cgn2_6/prodata/1/paa/US104_COMB.pep.*
31: /cgn2_6/prodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	518	100.0	518	25	US-09-991-681-27
2	518	100.0	1770	1	PCT-US03-01943-44

3	518	100.0	1770	27	US-10-144-198-44	Sequence 44, Appl
4	518 <th>100.0</th> <td>1807</td> <td>1</td> <td>PCT-US01-08631-40087</td> <td>Sequence 40087, A</td>	100.0	1807	1	PCT-US01-08631-40087	Sequence 40087, A
5	518 <th>100.0</th> <td>1839</td> <td>1</td> <td>PCT-US01-42950-495</td> <td>Sequence 495, App</td>	100.0	1839	1	PCT-US01-42950-495	Sequence 495, App
6	518 <th>100.0</th> <td>1839</td> <td>30</td> <td>US-10-416-993-495</td> <td>Sequence 495, App</td>	100.0	1839	30	US-10-416-993-495	Sequence 495, App
7	518 <th>100.0</th> <td>1872</td> <td>1</td> <td>PCT-US03-04508-32</td> <td>Sequence 32, Appl</td>	100.0	1872	1	PCT-US03-04508-32	Sequence 32, Appl
8	518 <th>100.0</th> <td>2221</td> <td>1</td> <td>PCT-US03-01943-30</td> <td>Sequence 30, Appl</td>	100.0	2221	1	PCT-US03-01943-30	Sequence 30, Appl
9	518 <th>100.0</th> <td>2221</td> <td>27</td> <td>US-10-144-198-30</td> <td>Sequence 30, Appl</td>	100.0	2221	27	US-10-144-198-30	Sequence 30, Appl
10	467 <th>90.2</th> <td>467</td> <td>12</td> <td>US-08-842-385-6</td> <td>Sequence 6, Appl</td>	90.2	467	12	US-08-842-385-6	Sequence 6, Appl
11	451 <th>87.1</th> <td>1982</td> <td>1</td> <td>PCT-US01-08631-40090</td> <td>Sequence 40090, A</td>	87.1	1982	1	PCT-US01-08631-40090	Sequence 40090, A
12	192 <th>37.1</th> <td>192</td> <td>20</td> <td>US-09-623-791-87</td> <td>Sequence 87, Appl</td>	37.1	192	20	US-09-623-791-87	Sequence 87, Appl
13	192 <th>37.1</th> <td>192</td> <td>20</td> <td>US-09-623-791A-87</td> <td>Sequence 87, Appl</td>	37.1	192	20	US-09-623-791A-87	Sequence 87, Appl
14	192 <th>37.1</th> <td>192</td> <td>27</td> <td>US-10-131-487A-87</td> <td>Sequence 87, Appl</td>	37.1	192	27	US-10-131-487A-87	Sequence 87, Appl
15	78 <th>15.1</th> <td>141</td> <td>28</td> <td>US-10-221-279-7783</td> <td>Sequence 7783, Ap</td>	15.1	141	28	US-10-221-279-7783	Sequence 7783, Ap
16	49 <th>9.5</th> <td>49</td> <td>12</td> <td>US-08-842-385-10</td> <td>Sequence 10, Appl</td>	9.5	49	12	US-08-842-385-10	Sequence 10, Appl
17	49 <th>9.5</th> <td>49</td> <td>25</td> <td>US-09-991-681-31</td> <td>Sequence 31, Appl</td>	9.5	49	25	US-09-991-681-31	Sequence 31, Appl
18	43 <th>8.3</th> <td>50</td> <td>1</td> <td>PCT-US01-08631-40088</td> <td>Sequence 40088, A</td>	8.3	50	1	PCT-US01-08631-40088	Sequence 40088, A
19	41 <th>7.9</th> <td>41</td> <td>12</td> <td>US-08-842-385-7</td> <td>Sequence 7, Appl</td>	7.9	41	12	US-08-842-385-7	Sequence 7, Appl
20	41 <th>7.9</th> <td>41</td> <td>25</td> <td>US-09-991-681-28</td> <td>Sequence 28, Appl</td>	7.9	41	25	US-09-991-681-28	Sequence 28, Appl
21	40 <th>7.7</th> <td>40</td> <td>12</td> <td>US-08-842-385-9</td> <td>Sequence 9, Appl</td>	7.7	40	12	US-08-842-385-9	Sequence 9, Appl
22	40 <th>7.7</th> <td>40</td> <td>25</td> <td>US-09-991-681-30</td> <td>Sequence 30, Appl</td>	7.7	40	25	US-09-991-681-30	Sequence 30, Appl
23	35 <th>6.8</th> <td>35</td> <td>12</td> <td>US-08-842-385-8</td> <td>Sequence 8, Appl</td>	6.8	35	12	US-08-842-385-8	Sequence 8, Appl
24	35 <th>6.8</th> <td>35</td> <td>25</td> <td>US-09-991-681-29</td> <td>Sequence 29, Appl</td>	6.8	35	25	US-09-991-681-29	Sequence 29, Appl
25	9 <th>1.7</th> <td>108</td> <td>22</td> <td>US-09-791-537-71315</td> <td>Sequence 71315, A</td>	1.7	108	22	US-09-791-537-71315	Sequence 71315, A
26	9 <th>1.7</th> <td>216</td> <td>19</td> <td>US-09-583-110-5166</td> <td>Sequence 5166, Ap</td>	1.7	216	19	US-09-583-110-5166	Sequence 5166, Ap
27	9 <th>1.7</th> <td>222</td> <td>15</td> <td>US-09-107-433-4471</td> <td>Sequence 4471, Ap</td>	1.7	222	15	US-09-107-433-4471	Sequence 4471, Ap
28	8 <th>1.5</th> <td>32</td> <td>1</td> <td>PCT-US01-00663-28863</td> <td>Sequence 28863, A</td>	1.5	32	1	PCT-US01-00663-28863	Sequence 28863, A
29	8 <th>1.5</th> <td>32</td> <td>23</td> <td>US-09-864-761-35805</td> <td>Sequence 35805, A</td>	1.5	32	23	US-09-864-761-35805	Sequence 35805, A
30	8 <th>1.5</th> <td>32</td> <td>27</td> <td>US-10-182-993-28015</td> <td>Sequence 28015, A</td>	1.5	32	27	US-10-182-993-28015	Sequence 28015, A
31	8 <th>1.5</th> <td>32</td> <td>27</td> <td>US-10-182-995-22277</td> <td>Sequence 22277, A</td>	1.5	32	27	US-10-182-995-22277	Sequence 22277, A
32	8 <th>1.5</th> <td>32</td> <td>27</td> <td>US-10-182-997-20929</td> <td>Sequence 20929, A</td>	1.5	32	27	US-10-182-997-20929	Sequence 20929, A
33	8 <th>1.5</th> <td>32</td> <td>27</td> <td>US-10-182-998-12570</td> <td>Sequence 12570, A</td>	1.5	32	27	US-10-182-998-12570	Sequence 12570, A
34	8 <th>1.5</th> <td>32</td> <td>28</td> <td>US-10-203-133-28586</td> <td>Sequence 28586, A</td>	1.5	32	28	US-10-203-133-28586	Sequence 28586, A
35	8 <th>1.5</th> <td>32</td> <td>28</td> <td>US-10-203-135-27486</td> <td>Sequence 27486, A</td>	1.5	32	28	US-10-203-135-27486	Sequence 27486, A
36	8 <th>1.5</th> <td>32</td> <td>28</td> <td>US-10-203-136-28586</td> <td>Sequence 28586, A</td>	1.5	32	28	US-10-203-136-28586	Sequence 28586, A
37	8 <th>1.5</th> <td>32</td> <td>28</td> <td>US-10-203-137-28863</td> <td>Sequence 28863, A</td>	1.5	32	28	US-10-203-137-28863	Sequence 28863, A
38	8 <th>1.5</th> <td>32</td> <td>28</td> <td>US-10-203-138-12879</td> <td>Sequence 12879, A</td>	1.5	32	28	US-10-203-138-12879	Sequence 12879, A
39	8 <th>1.5</th> <td>32</td> <td>28</td> <td>US-10-203-139A-12879</td> <td>Sequence 12879, A</td>	1.5	32	28	US-10-203-139A-12879	Sequence 12879, A
40	8 <th>1.5</th> <td>32</td> <td>28</td> <td>US-10-203-139-27724</td> <td>Sequence 27724, A</td>	1.5	32	28	US-10-203-139-27724	Sequence 27724, A
41	8 <th>1.5</th> <td>32</td> <td>31</td> <td>US-60-236-359-17560</td> <td>Sequence 17560, A</td>	1.5	32	31	US-60-236-359-17560	Sequence 17560, A
42	8 <th>1.5</th> <td>52</td> <td>30</td> <td>US-10-424-599-265455</td> <td>Sequence 265455, A</td>	1.5	52	30	US-10-424-599-265455	Sequence 265455, A
43	8 <th>1.5</th> <td>58</td> <td>30</td> <td>US-10-424-599-265456</td> <td>Sequence 265456, A</td>	1.5	58	30	US-10-424-599-265456	Sequence 265456, A
44	8 <th>1.5</th> <td>101</td> <td>30</td> <td>US-10-437-963-203739</td> <td>Sequence 203739, A</td>	1.5	101	30	US-10-437-963-203739	Sequence 203739, A
45	8 <th>1.5</th> <td>121</td> <td>20</td> <td>US-09-629-469A-18627</td> <td>Sequence 18627, A</td>	1.5	121	20	US-09-629-469A-18627	Sequence 18627, A

ALIGNMENTS

RESULT 1
US-09-991-681-27
Sequence 27, Application US/0991681
GENERAL INFORMATION:
APPLICANT: BILING-MEDEL, PATRICIA
COHEN, MAURICE
COLPITTS, TRACEY L.
FRIEDMAN, PAULA N.
GORDON, JULIAN
GRANADOS, EDWARD N.
HODGES, STEVEN C.
KLASS, MICHAEL R.
KRATOCHVIL, JON D.
ROBERTS-RAPP, LISA
TITLE OF INVENTION: FOR DETECTING AND METHODS USEFUL
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESS: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/991,681
FILING DATE: 26-Nov-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/065,383
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6084.US.PI
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 518 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: None
SEQUENCE DESCRIPTION: SEQ ID NO: 27:
US-09-991-681-27

Query Match 100.0%; Score 518; DB 25; Length 518;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIRMAQOVFMDLTQCSPTNNFHDHQSCLILPEPDEKNGHTKKSVSFRIVLSIL 60
DB 1 RIRMAQOVFMDLTQCSPTNNFHDHQSCLILPEPDEKNGHTKKSVSFRIVLSIL 60

QY 61 SHQVLQNLVYLLEEFYKGPSPEBEKTIQVPEAKLAGFLRYSIMONLAVIFDILLDSYR 120
DB 61 SHQVLQNLVYLLEEFYKGPSPEBEKTIQVPEAKLAGFLRYSIMONLAVIFDILLDSYR 120

QY 121 TAREFDTSPGLKCLLKTVSGIGGAANLYROSAMSFNIYFHALVCAVLTNOETITAEQYKK 180
DB 121 TAREFDTSPGLKCLLKTVSGIGGAANLYROSAMSFNIYFHALVCAVLTNOETITAEQYKK 180

QY 121 TAREFDTSPGLKCLLKTVSGIGGAANLYROSAMSFNIYFHALVCAVLTNOETITAEQYKK 180
DB 121 TAREFDTSPGLKCLLKTVSGIGGAANLYROSAMSFNIYFHALVCAVLTNOETITAEQYKK 180

QY 181 VLFEDDERSTDSOQCSSEDEDIFEEITAQVSPRGKEKRWARMPLLSVQVSNADVMW 240
DB 181 VLFEDDERSTDSOQCSSEDEDIFEEITAQVSPRGKEKRWARMPLLSVQVSNADVMW 240

QY 181 VLFEDDERSTDSOQCSSEDEDIFEEITAQVSPRGKEKRWARMPLLSVQVSNADVMW 240
DB 181 VLFEDDERSTDSOQCSSEDEDIFEEITAQVSPRGKEKRWARMPLLSVQVSNADVMW 240

QY 241 LVKRLHKLQMLCNNTYQMHLDLNCMBEPPIFKGDFFILPSFQSSSTPTSTGFSGKE 300
DB 241 LVKRLHKLQMLCNNTYQMHLDLNCMBEPPIFKGDFFILPSFQSSSTPTSTGFSGKE 300

QY 301 TPSEDDRSQSRHNGESLSLAKAGGDLPLPSPKVEKKDPSRKKEWMENAGNKITYMAAD 360
DB 301 TPSEDDRSQSRHNGESLSLAKAGGDLPLPSPKVEKKDPSRKKEWMENAGNKITYMAAD 360

QY 301 TPSEDDRSQSRHNGESLSLAKAGGDLPLPSPKVEKKDPSRKKEWMENAGNKITYMAAD 360
DB 301 TPSEDDRSQSRHNGESLSLAKAGGDLPLPSPKVEKKDPSRKKEWMENAGNKITYMAAD 360

QY 361 KTISKLMTEYKRRQOHNLSAFPKVEYKKGEPGPGGDSPLIORPOHLMDOGOMRHS 420
DB 361 KTISKLMTEYKRRQOHNLSAFPKVEYKKGEPGPGGDSPLIORPOHLMDOGOMRHS 420

QY 421 FSAGPELLRQDRPRSSTGSSLSVSRDAEAQIOAMTNMVLTVLNOIQILPDQFTALQ 480
DB 421 FSAGPELLRQDRPRSSTGSSLSVSRDAEAQIOAMTNMVLTVLNOIQILPDQFTALQ 480

QY 481 PAVFPCISQLTCHVTDIRVROAVREMLGRVGRVYDIIV 518
DB 481 PAVFPCISQLTCHVTDIRVROAVREMLGRVGRVYDIIV 518

RESULT 2
PCT-US03-01943-44
Sequence 44, Application PC/TUS0301943

GENERAL INFORMATION:
APPLICANT: ORIGENE TECHNOLOGIES INC
TITLE OF INVENTION: CANCER GENES
FILE REFERENCE: 3U 9U 901 PCT
CURRENT APPLICATION NUMBER: PCT/US03/01943
PRIOR FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: US 10/054,935
PRIOR FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: US 60/356,130
PRIOR FILING DATE: 2002-02-14
PRIOR APPLICATION NUMBER: US 10/102,946
PRIOR FILING DATE: 2002-03-22
PRIOR APPLICATION NUMBER: US 10/117,229
PRIOR FILING DATE: 2002-04-08
PRIOR APPLICATION NUMBER: US 10/144,198
PRIOR FILING DATE: 2002-05-14
PRIOR APPLICATION NUMBER: US 10/197,824
PRIOR FILING DATE: 2002-07-19
NUMBER OF SEQ ID NOS: 102
SOFTWARE: PatentIn version 3.1
SEQ ID NO 44
LENGTH: 1770
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US03-01943-44

Query Match 100.0%; Score 518; DB 1; Length 1770;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIRMAQOVFMDLTQCSPTNNFHDHQSCLILPEPDEKNGHTKKSVSFRIVLSIL 60
DB 1253 RIRMAQOVFMDLTQCSPTNNFHDHQSCLILPEPDEKNGHTKKSVSFRIVLSIL 1312

QY 61 SHQVLQNLVYLLEEFYKGPSPEBEKTIQVPEAKLAGFLRYSIMONLAVIFDILLDSYR 120
DB 1313 SHQVLQNLVYLLEEFYKGPSPEBEKTIQVPEAKLAGFLRYSIMONLAVIFDILLDSYR 1372

QY 121 TAREFDTSPGLKCLLKTVSGIGGAANLYROSAMSFNIYFHALVCAVLTNOETITAEQYKK 180
DB 1373 TAREFDTSPGLKCLLKTVSGIGGAANLYROSAMSFNIYFHALVCAVLTNOETITAEQYKK 1432

QY 181 VLFEDDERSTDSOQCSSEDEDIFEEITAQVSPRGKEKRWARMPLLSVQVSNADVMW 240
DB 1433 VLFEDDERSTDSOQCSSEDEDIFEEITAQVSPRGKEKRWARMPLLSVQVSNADVMW 1492

QY 241 LVKRLHKLQMLCNNTYQMHLDLNCMBEPPIFKGDFFILPSFQSSSTPTSTGFSGKE 300
DB 1493 LVKRLHKLQMLCNNTYQMHLDLNCMBEPPIFKGDFFILPSFQSSSTPTSTGFSGKE 1552

QY 301 TPSEDDRSQSRHNGESLSLAKAGGDLPLPSPKVEKKDPSRKKEWMENAGNKITYMAAD 360
DB 1553 TPSEDDRSQSRHNGESLSLAKAGGDLPLPSPKVEKKDPSRKKEWMENAGNKITYMAAD 1612

QY 361 KTISKLMTEYKRRQOHNLSAFPKVEYKKGEPGPGGDSPLIORPOHLMDOGOMRHS 420
DB 1613 KTISKLMTEYKRRQOHNLSAFPKVEYKKGEPGPGGDSPLIORPOHLMDOGOMRHS 1672

QY 421 FSAGPELLRQDRPRSSTGSSLSVSRDAEAQIOAMTNMVLTVLNOIQILPDQFTALQ 480
DB 1673 FSAGPELLRQDRPRSSTGSSLSVSRDAEAQIOAMTNMVLTVLNOIQILPDQFTALQ 1732

QY 481 PAVFPCISQLTCHVTDIRVROAVREMLGRVGRVYDIIV 518
DB 1733 PAVFPCISQLTCHVTDIRVROAVREMLGRVGRVYDIIV 1770

RESULT 3
US-10-144-198-44
Sequence 44, Application US/10144198
GENERAL INFORMATION:
APPLICANT: Origene Technologies Inc
TITLE OF INVENTION: Regulated Prostate Cance Genes

```

RESULT 4
PCT-US01-08631-40087
Sequence 40087, Application PC/TUS0108631
GENERAL INFORMATION:
APPLICANT: Hyeeq, Inc
FILE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 21272-049
CURRENT APPLICATION NUMBER: PCT/US01/08631
CURRENT FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 09/540,217
PRIOR FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: 09/649,167
PRIOR FILING DATE: 2000-08-23
NUMBER OF SEQ ID NOS: 60736
SOFTWARE: Custom
SEQ ID NO 40087
LENGTH: 1807
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: DOMAIN

```

Query Match	100.0%	Score 518;	DB 1;	Length 1639;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 518;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	1	RIRANAQOVFMJLDTQCSPTPNPNFPAHOSCOLIETELPREKNGHTKKSVSRELVTSL	60
Db	1322	RIRANAQOVFMJLDTQCSPTPNPNFPAHOSCOLIETELPREKNGHTKKSVSRELVTSL	1387
Qy	61	SHOVLONLYDILLEEPAVKPSPGSEKTIQVEBAKIAGFLRYISMONLAVIFDLILDSYR	120
Db	1382	SHOVLONLYDILLEEPAVKPSPGSEKTIQVEBAKIAGFLRYISMONLAVIFDLILDSYR	1443
Qy	121	TAREFDTSPLGKCLLKVKYSGIGCANLROSMSNFIYFHALVGVLTNQEFTTAEOYKX	180
Db	1442	TAREFDTSPLGKCLLKVKYSGIGCANLROSMSNFIYFHALVGVLTNQEFTTAEOYKX	1501
Qy	181	VLFEDEDERSTBSSQOCCSDEDEIPEETAQVSPPRGEXKQWAPRMPLLSVQVSNADWW	240
Db	1502	VLFEDEDERSTBSSQOCCSDEDEIPEETAQVSPPRGEXKQWAPRMPLLSVQVSNADWW	1561
Qy	241	LVKRLHKLAMELCNNYIQOMHLLENKMEBPPLEFKDDPFFILPSPOSESTPTGSGSGE	300
Db	1562	LVKRLHKLAMELCNNYIQOMHLLENKMEBPPLEFKDDPFFILPSPOSESTPTGSGSGE	1621
Qy	301	TPSEDDRSQSRHEHMGESLSLKAQGGDLLPSPKPYEKDPSRKKEWENAGNKIYTMAD	360
Db	1622	TPSEDDRSQSRHEHMGESLSLKAQGGDLLPSPKPYEKDPSRKKEWENAGNKIYTMAD	1681
Qy	361	KTISKLMEYKRRKQONHLSAPPEKVYXKEKCEPLGPRGQDSPLLQRPOLMDQOMKHS	420
Db	1662	KTISKLMEYKRRKQONHLSAPPEKVYXKEKCEPLGPRGQDSPLLQRPOLMDQOMKHS	1743
Qy	421	FSAGBELLRODKRPPRSGSTGSLSVSVDAEAOIQAMTNMVLTVLNOIQIILPDQFTALQ	480
Db	1742	FSAGBELLRODKRPPRSGSTGSLSVSVDAEAOIQAMTNMVLTVLNOIQIILPDQFTALQ	1801
Qy	481	PAVPFCISQLTCHVTDIRVQAVRMLRGVGVYIIV	518
Db	1802	PAVPFCISQLTCHVTDIRVQAVRMLRGVGVYIIV	1839

```

RESULT 6
US-10-416-993-495
; Sequence 495: Application US/10416993
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-096
; CURRENT APPLICATION NUMBER: US/10/416,993
; CURRENT FILING DATE: 2003-11-16
; PRIOR APPLICATION NUMBER: 09/714,936
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 682
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 495
; LENGTH: 1839
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-416-993-495

Query Match          100.0%; Score 518; DB 30; Length 1839;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY      1 RIRAAQGVFMLDTCSSPKTPNNFPHAAQSCQIIELPDEKNGHTKKSVSFREIVSL 60
Db      1322 RIRAAQGVFMLDTCSSPKTPNNFPHAAQSCQIIELPDEKNGHTKKSVSFREIVSL 1382
          |||
QY      61 SHOVLTQWLYDILLEEFVKGSPGSEKTIQVPEATLQAGLRYISQNLAVIFDILLDSYR 120
Db      1382 SHOVLTQWLYDILLEEFVKGSPGSEKTIQVPEATLQAGLRYISQNLAVIFDILLDSYR 1443
          |||
QY      121 TAREFDTSPGKCLTKLKVSGIGAGANLYRQASMSFNIFYHALVCAVLTNQETITAEQYK 180
Db      1442 TAREFDTSPGKCLTKLKVSGIGAGANLYRQASMSFNIFYHALVCAVLTNQETITAEQYK 1501
          |||
QY      181 VLFEDDEASTSSQQCCSEDEDIIEFTYQAVSPPEKREKRQWRAMPLLISVQDPVSADAWM 240

```

Db	1502	VLFEDDERSTDS	SQOCCSEDD	IFEBTR	QUSP	PGCKEKQ	RAMPL	PLISV	PVSNAB	VYM	1566
Qy	241	LVKLAHLKCMEL	CNNYI	IQMHL	DLENCK	EEPTI	FKGDPP	FLIP	SFOSBS	STPSTG	SGSKE 300
Db	1562	LVKLAHLKCMEL	CNNYI	IQMHL	DLENCK	EEPTI	FKGDPP	FLIP	SFOSBS	STPSTG	SGSKE 1622
Qy	301	TPSRDDRSQS	REHNGES	LSLKAGG	DLLP	SPSV	VEKKD	PSRKKE	WENNA	CNK1	YTMAD 360
Db	1622	TPSRDDRSQS	REHNGES	LSLKAGG	DLLP	SPSV	VEKKD	PSRKKE	WENNA	CNK1	YTMAD 1681
Qy	361	KTISKLTMT	EYKKRQ	QOHNLSA	FP	PEVYE	EKKGE	PLGR	QDS	PLJOR	PHILMDQOMRHS 420
Db	1682	KTISKLTMT	EYKKRQ	QOHNLSA	FP	PEVYE	EKKGE	PLGR	QDS	PLJOR	PHILMDQOMRHS 1741
Qy	421	FSAPPELLRQ	KRRSG	SGSSLS	VS	VRDA	MAQIQ	AMTN	NWLT	VLNO	QIILPDQFTALQ 480
Db	1742	FSAPPELLRQ	KRRSG	SGSSLS	VS	VRDA	MAQIQ	AMTN	NWLT	VLNO	QIILPDQFTALQ 1801
Qy	481	PAVEPCISOL	TCHVTDIR	VA	VBEM	IGR	VR	YVDI	IV	518	
Db	1802	PAVEPCISOL	TCHVTDIR	VA	VBEM	IGR	VR	YVDI	IV	1839	

```

RESULT 7
PCT-US03-04508-32
; Sequence 32, Application PC/TUS0304508
; GENERAL INFORMATION:
; APPLICANT: IDEC PHARMACEUTICALS
; APPLICANT: GATELY, DENNIS
; TITLE OF INVENTION: PROSTATE SPECIFIC GENES AND THE USE THEREOF IN DESIGN
; TITLE OF INVENTION: OR THERAPEUTICS
; FILE REFERENCE: 037003/0301985
; CURRENT APPLICATION NUMBER: PCT/US03/04508
; CURRENT FILING DATE: 2003-02-19
; PRIOR APPLICATION NUMBER: 60/357,140
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: 60/396,082
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 60/386,759
; PRIOR FILING DATE: 2002-06-10
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 1872
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-04508-32

```

	Query Match	100.0%;	Score 518;	DB 1;	Length 1872;
	Best Local Similarity	100.0%;	Pred. No. 0;		
	Matches 518;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	RIRAAQOVFEMLDTPQCSPKTPPNDFHAQSCOLIIELPDPDEKPNCHTKXSVSFREIVSVLL	60		
Db	1355	RIRAAQOVFEMLDTPQCSPKTPPNDFHAQSCOLIIELPDPDEKPNCHTKXSVSFREIVSVLL	1414		
Qy	61	SHOVLQNLVDILLEEYVKGSPSEEEKITQVPEAKLGLFLYISMONLAVIFDILLDSYR	120		
Db	1415	SHOVLQNLVDILLEEYVKGSPSEEEKITQVPEAKLGLFLYISMONLAVIFDILLDSYR	1474		
Qy	121	TAREFDTSFGIKCLLKLYSGIGGAANTYROSAMSFNITYFHALVCAVLTNDETTIAEYOYKK	180		
Db	1475	TAREFDTSFGIKCLLKLYSGIGGAANTYROSAMSFNITYFHALVCAVLTNDETTIAEYOYKK	1534		
Qy	181	VLFEEDDERSTDSQOCCSSDEDIIEETAAQVSPPRGKEKROQRAMPPLLSVQPVANADVMV	240		
Db	1535	VLFEEDDERSTDSQOCCSSDEDIIEETAAQVSPPRGKEKROQRAMPPLLSVQPVANADVMV	1594		
Qy	241	LYVKRLHLKCMELCNNTYQMHLDLENCEBEPPIFGDGPPFLIPSQOSSSTSTSTGFGSKE	300		
Db	1595	LYVKRLHLKCMELCNNTYQMHLDLENCEBEPPIFGDGPPFLIPSQOSSSTSTSTGFGSKE	1654		

```
Qy 301 TPSEDDRSQSRHNGESLSLAKGGDLLPPSPVKEKDDPSRKKEWMENAGNKIYTMAAD 360
Db 1655 TPSEDDRSQSRHNGESLSLAKGGDLLPPSPVKEKDDPSRKKEWMENAGNKIYTMAAD 1714
Qy 361 KTISKMTYKKRQKQHNLSAFPKEVYKKEGEBLGRGDSPLQRPQHLMDQGWRHS 420
Db 1715 KTISKMTYKKRQKQHNLSAFPKEVYKKEGEBLGRGDSPLQRPQHLMDQGWRHS 1774
Qy 421 FSAGPELLRQDKRPRSGSTSSLSVSVRADEAOIQAWTNVLTVLNQIQLPDTFTALQ 480
Db 1775 FSAGPELLRQDKRPRSGSTSSLSVSVRADEAOIQAWTNVLTVLNQIQLPDTFTALQ 1834
Qy 481 PAVFPCISQLTCHVTDIRVROAVREMLGRVGRVYDIIV 518
Db 1835 PAVFPCISQLTCHVTDIRVROAVREMLGRVGRVYDIIV 1872

RESULT 8
PCT-US03-01943-30
; Sequence 30, Application PC/TUS0301943
; GENERAL INFORMATION:
; APPLICANT: ORIGENE TECHNOLOGIES INC
; TITLE OF INVENTION: CANCER GENES
; FILE REFERENCE: 3U 9U 901 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/01943
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: US 10/054,935
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: US 60/356,130
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 10/102,946
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: US 10/117,229
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: US 10/144,198
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: US 10/197,824
; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30
; LENGTH: 2221
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-01943-30

Query Match 100.0%; Score 518; DB 1; Length 2221;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RIRAMAQOVFMLDTQCSPTKPNPFDAQSCQLIILEPDEKPNGHTTKSVSFREIVVSL 60
Db 1704 RIRAMAQOVFMLDTQCSPTKPNPFDAQSCQLIILEPDEKPNGHTTKSVSFREIVVSL 1763
Qy 61 SHQVLLQNLVDILLEEFVKGPSPEBEKTIQVPEAKLAGFLRYISMOMLAVIFDLLDSYR 120
Db 1764 SHQVLLQNLVDILLEEFVKGPSPEBEKTIQVPEAKLAGFLRYISMOMLAVIFDLLDSYR 1823
Qy 121 TAREFDTSPGLKCLIKKVSIGGANLYROSAMSFNIYFHALVCAVLTNOETITAEQYKK 180
Db 1824 TAREFDTSPGLKCLIKKVSIGGANLYROSAMSFNIYFHALVCAVLTNOETITAEQYKK 1883
Qy 181 VLFEDDERSTDSQOCCSSEDEDIIEETAOVSPPRGKEKQWRAMPPLSVQPVSNADVMW 240
Db 1884 VLFEDDERSTDSQOCCSSEDEDIIEETAOVSPPRGKEKQWRAMPPLSVQPVSNADVMW 1943
Qy 241 LVKRLHKLCHMELCNNTYQMHLDLENCEBEPPIFKGDPFFILPSFQSSSTPSTGFSGKE 300
Db 1944 LVKRLHKLCHMELCNNTYQMHLDLENCEBEPPIFKGDPFFILPSFQSSSTPSTGFSGKE 2003
Qy 301 TPSEDDRSQSRHNGESLSLAKGGDLLPPSPVKEKDDPSRKKEWMENAGNKIYTMAAD 360
Db 2004 TPSEDDRSQSRHNGESLSLAKGGDLLPPSPVKEKDDPSRKKEWMENAGNKIYTMAAD 2063
```

```
RESULT 9
US-10-144-198-30
; Sequence 30, Application US/10144198
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies Inc
; TITLE OF INVENTION: Regulated Prostate Cance Genes
; FILE REFERENCE: 9U 105 R1
; CURRENT APPLICATION NUMBER: US/10/144,198
; CURRENT FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 30
; LENGTH: 2221
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-198-30
```

Query Match 100.0%; Score 518; DB 27; Length 2221;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 RIRAMAQOVFMLDTQCSPTKPNPFDAQSCQLIILEPDEKPNGHTTKSVSFREIVVSL 60
Db 1704 RIRAMAQOVFMLDTQCSPTKPNPFDAQSCQLIILEPDEKPNGHTTKSVSFREIVVSL 1763
Qy 61 SHQVLLQNLVDILLEEFVKGPSPEBEKTIQVPEAKLAGFLRYISMOMLAVIFDLLDSYR 120
Db 1764 SHQVLLQNLVDILLEEFVKGPSPEBEKTIQVPEAKLAGFLRYISMOMLAVIFDLLDSYR 1823
Qy 121 TAREFDTSPGLKCLIKKVSIGGANLYROSAMSFNIYFHALVCAVLTNOETITAEQYKK 180
Db 1824 TAREFDTSPGLKCLIKKVSIGGANLYROSAMSFNIYFHALVCAVLTNOETITAEQYKK 1883
Qy 181 VLFEDDERSTDSQOCCSSEDEDIIEETAOVSPPRGKEKQWRAMPPLSVQPVSNADVMW 240
Db 1884 VLFEDDERSTDSQOCCSSEDEDIIEETAOVSPPRGKEKQWRAMPPLSVQPVSNADVMW 1943
Qy 241 LVKRLHKLCHMELCNNTYQMHLDLENCEBEPPIFKGDPFFILPSFQSSSTPSTGFSGKE 300
Db 1944 LVKRLHKLCHMELCNNTYQMHLDLENCEBEPPIFKGDPFFILPSFQSSSTPSTGFSGKE 2003
Qy 301 TPSEDDRSQSRHNGESLSLAKGGDLLPPSPVKEKDDPSRKKEWMENAGNKIYTMAAD 360
Db 2004 TPSEDDRSQSRHNGESLSLAKGGDLLPPSPVKEKDDPSRKKEWMENAGNKIYTMAAD 2063
Qy 361 KTISKMTYKKRQKQHNLSAFPKEVYKKEGEBLGRGDSPLQRPQHLMDQGWRHS 420
Db 2064 KTISKMTYKKRQKQHNLSAFPKEVYKKEGEBLGRGDSPLQRPQHLMDQGWRHS 2123
Qy 421 FSAGPELLRQDKRPRSGSTSSLSVSVRADEAOIQAWTNVLTVLNQIQLPDTFTALQ 480
Db 2124 FSAGPELLRQDKRPRSGSTSSLSVSVRADEAOIQAWTNVLTVLNQIQLPDTFTALQ 2183
Qy 481 PAVFPCISQLTCHVTDIRVROAVREMLGRVGRVYDIIV 518
Db 2184 PAVFPCISQLTCHVTDIRVROAVREMLGRVGRVYDIIV 2221
```

RESULT 10

US-08-842-385-6
; Sequence 6, Application US/08842385
; GENERAL INFORMATION:
; APPLICANT: Russel, John
; APPLICANT: Colpitts, Tracey
; TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
; TITLE OF INVENTION: FOR DETECTING DISEASE OF THE PROSTATE
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/842.385
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Porembski, Priscilla E
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 6084.US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847/937-6365
; TELEFAX: 847/938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: None
; US-08-842-385-6

Query Match 90.2%; Score 467; DB 12; Length 467;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 467; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 FREIVSLSHQVLLQNLVDLLEFVKGSPGSEKTIQVPAKLAGFLRYISMONTLAVI 111
DB 1 FREIVSLSHQVLLQNLVDLLEFVKGSPGSEKTIQVPAKLAGFLRYISMONTLAVI 60
QY 112 FDLILDSYRTAREPFTSPGLKCLLKKVSGIGGANLYRQSAASFNIYFHALVCALVTNOE 171
DB 61 FDLILDSYRTAREPFTSPGLKCLLKKVSGIGGANLYRQSAASFNIYFHALVCALVTNOE 120
QY 172 TITAEQVKVLFEDDERSTDSQQCSSEDEDIFFETAAVSPRGKEKQWRAMPPLLSVQ 231
DB 121 TITAEQVKVLFEDDERSTDSQQCSSEDEDIFFETAAVSPRGKEKQWRAMPPLLSVQ 180
QY 232 PVSNADWMLVLRHLKLCMELCNNTYQMHLDLENCEBPPIFKGDPFLLPSFQSSSTP 291
DB 181 PVSNADWMLVLRHLKLCMELCNNTYQMHLDLENCEBPPIFKGDPFLLPSFQSSSTP 240
QY 292 STGGSGKETPSEDROSREHMGESLSLKAGGGLLLPSPKVKKQPSRKKEWENAG 351
DB 241 STGGSGKETPSEDROSREHMGESLSLKAGGGLLLPSPKVKKQPSRKKEWENAG 300
QY 352 NKIYMAADKTIISKLMTEYKRRKQOHNLAPPKVEKVEKKGEPPLGPRGQDSPLLQRPQHL 411
DB 301 NKIYMAADKTIISKLMTEYKRRKQOHNLAPPKVEKVEKKGEPPLGPRGQDSPLLQRPQHL 360
QY 412 MDQGMRSFSAGPBLAQDRKPRSGSTGSSLSVSVRAEAQIQMTNMTLVLTNIOIQL 471

DB 361 MDQGMRSFSAGPBLAQDRKPRSGSTGSSLSVSVRAEAQIQMTNMTLVLTNIOIQL 420
QY 472 PDQFTALQPAVFPICISQLTCHVTDIRVROAVREMLGRVGVYDIIV 518
DB 421 PDQFTALQPAVFPICISQLTCHVTDIRVROAVREMLGRVGVYDIIV 467

RESULT 11
PCT-US01-08631-40090
; Sequence 40090, Application PC/TUS0108631
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-049
; CURRENT APPLICATION NUMBER: PCT/US01/08631
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 40090
; LENGTH: 1982
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (11)-(25)
; OTHER INFORMATION: Fusion glycoprotein F0 domain identified by eMATRIX.
; OTHER INFORMATION: accession number PF00523D, p-value=7.188e-10, raw score of 11.39
; NAME/KEY: DOMAIN
; LOCATION: (1065)-(1074)
; OTHER INFORMATION: Helper component proteinase domain identified by Pfam,
; OTHER INFORMATION: accession name Peptidase_C6, E-value=0.0056, Pfam score of 7.9
PCT-US01-08631-40090

Query Match 87.1%; Score 451; DB 1; Length 1982;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 451; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 REIVSLSHQVLLQNLVDLLEFVKGSPGSEKTIQVPAKLAGFLRYISMONTLAVI 112
DB 1456 REIVSLSHQVLLQNLVDLLEFVKGSPGSEKTIQVPAKLAGFLRYISMONTLAVI 1515
QY 113 DILDSYRTAREPFTSPGLKCLLKKVSGIGGANLYRQSAASFNIYFHALVCALVTNOE 172
DB 1516 DILDSYRTAREPFTSPGLKCLLKKVSGIGGANLYRQSAASFNIYFHALVCALVTNOE 1575
QY 173 ITAEQVKVLFEDDERSTDSQQCSSEDEDIFFETAAVSPRGKEKQWRAMPPLLSVQ 232
DB 1576 ITAEQVKVLFEDDERSTDSQQCSSEDEDIFFETAAVSPRGKEKQWRAMPPLLSVQ 1635
QY 233 VSNADWMLVLRHLKLCMELCNNTYQMHLDLENCEBPPIFKGDPFLLPSFQSSSTP 292
DB 1636 VSNADWMLVLRHLKLCMELCNNTYQMHLDLENCEBPPIFKGDPFLLPSFQSSSTP 1695
QY 293 TGGSGKETPSEDROSREHMGESLSLKAGGGLLLPSPKVKKQPSRKKEWENAG 352
DB 1696 TGGSGKETPSEDROSREHMGESLSLKAGGGLLLPSPKVKKQPSRKKEWENAG 1755
QY 353 KIYMAADKTIISKLMTEYKRRKQOHNLAPPKVEKVEKKGEPPLGPRGQDSPLLQRPQHL 412
DB 1756 KIYMAADKTIISKLMTEYKRRKQOHNLAPPKVEKVEKKGEPPLGPRGQDSPLLQRPQHL 1815
QY 413 DQGMRSFSAGPBLAQDRKPRSGSTGSSLSVSVRAEAQIQMTNMTLVLTNIOIQL 472
DB 1816 DQGMRSFSAGPBLAQDRKPRSGSTGSSLSVSVRAEAQIQMTNMTLVLTNIOIQL 1875
QY 473 DQFTALQPAVFPICISQLTCHVTDIRVROAV 503
DB 1876 DQFTALQPAVFPICISQLTCHVTDIRVROAV 1906

RESULT 12
US-09-623-791-87
Sequence 87, Application US/09623791
GENERAL INFORMATION:
APPLICANT: SPECHT, THOMAS
APPLICANT: HINZMANN, BERND
APPLICANT: SCHMITT, ARMIN
APPLICANT: PILARSKY, CHRISTIAN
APPLICANT: DAHL, EDGAR
APPLICANT: ROSENTHAL, ANDRE
TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM PROSTATE TUMOR TISSUE
FILE REFERENCE: ALBRE 11
CURRENT APPLICATION NUMBER: US/09/623,791
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: PCT/DE99/00721
PRIOR FILING DATE: 1999-03-09
NUMBER OF SEQ ID NOS: 201
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 87
LENGTH: 192
TYPE: PRT
ORGANISM: Homo sapiens
US-09-623-791-87

Query Match 37.1%; Score 192; DB 20; Length 192;
Best Local Similarity 100.0%; Pred. No. 1.2e-184;
Matches 192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 MAQOVFMDTQCSPKTPNNPDHQAOSCOLITELPDEKPNKHTKKSVSFREIVVSLSHOV 64
DB 1 MAQOVFMDTQCSPKTPNNPDHQAOSCOLITELPDEKPNKHTKKSVSFREIVVSLSHOV 60
QY 65 LQNLVYDILLEEFVKGPSPEEKTIQVPEAKLAGFLRYSIMQNLAVIFDILLDSYRTARE 124
DB 61 LQNLVYDILLEEFVKGPSPEEKTIQVPEAKLAGFLRYSIMQNLAVIFDILLDSYRTARE 120
QY 125 FDTSPGKCLLKXVSGIGGANLYROSAMSFNIFHALVCAVLTNOETTITAEQVKVLF 184
DB 121 FDTSPGKCLLKXVSGIGGANLYROSAMSFNIFHALVCAVLTNOETTITAEQVKVLF 180
QY 185 DDERSTDSQOC 196
DB 181 DDERSTDSQOC 192

RESULT 13
US-09-623-791A-87
Sequence 87, Application US/09623791A
GENERAL INFORMATION:
APPLICANT: SPECHT, THOMAS
APPLICANT: HINZMANN, BERND
APPLICANT: SCHMITT, ARMIN
APPLICANT: PILARSKY, CHRISTIAN
APPLICANT: DAHL, EDGAR
APPLICANT: ROSENTHAL, ANDRE
TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM PROSTATE TUMOR TISSUE
FILE REFERENCE: ALBRE 11
CURRENT APPLICATION NUMBER: US/09/623,791A
CURRENT FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: PCT/DE99/00721
PRIOR FILING DATE: 1999-03-09
NUMBER OF SEQ ID NOS: 201
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 87
LENGTH: 192
TYPE: PRT
ORGANISM: Homo sapiens
US-09-623-791A-87

Query Match 37.1%; Score 192; DB 20; Length 192;
Best Local Similarity 100.0%; Pred. No. 1.2e-184;

Matches 192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5 MAQOVFMDTQCSPKTPNNPDHQAOSCOLITELPDEKPNKHTKKSVSFREIVVSLSHOV 64
DB 1 MAQOVFMDTQCSPKTPNNPDHQAOSCOLITELPDEKPNKHTKKSVSFREIVVSLSHOV 60
QY 65 LQNLVYDILLEEFVKGPSPEEKTIQVPEAKLAGFLRYSIMQNLAVIFDILLDSYRTARE 124
DB 61 LQNLVYDILLEEFVKGPSPEEKTIQVPEAKLAGFLRYSIMQNLAVIFDILLDSYRTARE 120
QY 125 FDTSPGKCLLKXVSGIGGANLYROSAMSFNIFHALVCAVLTNOETTITAEQVKVLF 184
DB 121 FDTSPGKCLLKXVSGIGGANLYROSAMSFNIFHALVCAVLTNOETTITAEQVKVLF 180
QY 185 DDERSTDSQOC 196
DB 181 DDERSTDSQOC 192

RESULT 14
US-10-131-487A-87
Sequence 87, Application US/10131487A
GENERAL INFORMATION:
APPLICANT: SPECHT, THOMAS
APPLICANT: HINZMANN, BERND
APPLICANT: SCHMITT, ARMIN
APPLICANT: PILARSKY, CHRISTIAN
APPLICANT: DAHL, EDGAR
APPLICANT: ROSENTHAL, ANDRE
TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM PROSTATE TUMOR TISSUE
FILE REFERENCE: ALBRE 11
CURRENT APPLICATION NUMBER: US/10/131,487A
CURRENT FILING DATE: 2002-04-25
PRIOR APPLICATION NUMBER: US/09/623,791A
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: PCT/DE99/00721
PRIOR FILING DATE: 1999-03-09
NUMBER OF SEQ ID NOS: 201
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 87
LENGTH: 192
TYPE: PRT
ORGANISM: Homo sapiens
US-10-131-487A-87

Query Match 37.1%; Score 192; DB 27; Length 192;
Best Local Similarity 100.0%; Pred. No. 1.2e-184;
Matches 192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 MAQOVFMDTQCSPKTPNNPDHQAOSCOLITELPDEKPNKHTKKSVSFREIVVSLSHOV 64
DB 1 MAQOVFMDTQCSPKTPNNPDHQAOSCOLITELPDEKPNKHTKKSVSFREIVVSLSHOV 60
QY 65 LQNLVYDILLEEFVKGPSPEEKTIQVPEAKLAGFLRYSIMQNLAVIFDILLDSYRTARE 124
DB 61 LQNLVYDILLEEFVKGPSPEEKTIQVPEAKLAGFLRYSIMQNLAVIFDILLDSYRTARE 120
QY 125 FDTSPGKCLLKXVSGIGGANLYROSAMSFNIFHALVCAVLTNOETTITAEQVKVLF 184
DB 121 FDTSPGKCLLKXVSGIGGANLYROSAMSFNIFHALVCAVLTNOETTITAEQVKVLF 180
QY 185 DDERSTDSQOC 196
DB 181 DDERSTDSQOC 192

RESULT 15
US-10-221-279-7783
Sequence 7783, Application US/10221279
GENERAL INFORMATION:
APPLICANT: Hysag, Inc
TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
FILE REFERENCE: 21272-046

```

; CURRENT APPLICATION NUMBER: US/10/221,279
; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 12360
; SOFTWARE: Custom
; SEQ ID NO: 7783
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(141)
; OTHER INFORMATION: Xaa = any amino acid or nothing
US-10-221-279-7783

```

```

Query Match      15.1%; Score 78; DB 28; Length 141;
Best Local Similarity 100.0%; Pred. No. 1.8e-69;
Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      298 GKEPSEDDRSQSRHMGESLSLKAGGDLPLPSPKVEKKDPSRKKEWENAGNKIYTM 357
      |||||||
DB      62  GKEPSEDDRSQSRHMGESLSLKAGGDLPLPSPKVEKKDPSRKKEWENAGNKIYTM 121
QY      358 AADKTISKLTMTYKKRKQ 375
      |||||||
DB      122 AADKTISKLTMTYKKRKQ 139

```

Search completed: July 24, 2003, 12:45:53
 Job time : 782.413 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: July 24, 2003, 12:22:19 ; Search time 15.1684 Seconds
(without alignments)
140.227 Million cell updates/sec

Title: US-09-991-681-27

Perfect score: 518

Sequence: 1 RIRMAQOVFMDTQCSPKT.....VRQAVREMLGRVGRVYDIIV 518

Scoring table: OLIGO

Searched: 41799 seqs, 4106219 residues

Word size: 0

Total number of hits satisfying chosen parameters: 41799

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: Pending Patents_AA New:*

1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	518	100.0	1872	6	US-10-367-978-32
2	7	1.4	10	6	US-10-462-850-3170
3	7	1.4	151	6	US-10-273-573-8627
4	7	1.4	418	5	US-09-291-417D-99
5	7	1.4	609	7	US-60-478-196-3273
6	7	1.4	642	6	US-10-273-573-10795
7	7	1.4	1162	6	US-10-294-433-224
8	7	1.4	1246	6	US-10-273-573-10799
9	6	1.2	10	6	US-10-462-850-104
10	6	1.2	10	6	US-10-462-850-3172
11	6	1.2	10	6	US-10-462-850-3684
12	6	1.2	10	6	US-10-462-850-3685
13	6	1.2	23	6	US-10-369-060A-111
14	6	1.2	40	6	US-10-273-573-7491
15	6	1.2	97	6	US-10-273-573-10762
16	6	1.2	151	6	US-10-273-573-7932
17	6	1.2	151	6	US-10-273-573-8420
18	6	1.2	155	6	US-10-273-573-8240
19	6	1.2	166	6	US-10-273-573-9011
20	6	1.2	173	6	US-10-273-573-9548
21	6	1.2	183	6	US-10-243-739A-33
22	6	1.2	183	6	US-10-243-739A-38
23	6	1.2	183	6	US-10-243-739A-39
24	6	1.2	183	6	US-10-243-739A-42
25	6	1.2	183	6	US-10-243-739A-44
26	6	1.2	183	6	US-10-243-739A-45

27	6	1.2	183	6	US-10-243-739A-46	Sequence 46, Appl
28	6	1.2	183	6	US-10-243-739A-59	Sequence 59, Appl
29	6	1.2	183	6	US-10-243-739A-60	Sequence 60, Appl
30	6	1.2	183	6	US-10-243-739A-63	Sequence 63, Appl
31	6	1.2	183	6	US-10-243-739A-65	Sequence 65, Appl
32	6	1.2	183	6	US-10-243-739A-71	Sequence 71, Appl
33	6	1.2	183	6	US-10-465-811-24	Sequence 29, Appl
34	6	1.2	183	6	US-10-465-811-29	Sequence 24, Appl
35	6	1.2	183	6	US-10-465-811-30	Sequence 30, Appl
36	6	1.2	183	6	US-10-465-811-33	Sequence 33, Appl
37	6	1.2	183	6	US-10-465-811-35	Sequence 35, Appl
38	6	1.2	183	6	US-10-465-811-36	Sequence 36, Appl
39	6	1.2	183	6	US-10-465-811-37	Sequence 37, Appl
40	6	1.2	183	6	US-10-465-811-50	Sequence 50, Appl
41	6	1.2	183	6	US-10-465-811-51	Sequence 51, Appl
42	6	1.2	183	6	US-10-465-811-54	Sequence 54, Appl
43	6	1.2	183	6	US-10-465-811-56	Sequence 56, Appl
44	6	1.2	183	6	US-10-465-811-62	Sequence 62, Appl
45	6	1.2	183	7	US-60-470-920-40	Sequence 40, Appl

ALIGNMENTS

RESULT 1
US-10-367-978-32 Application US/10367978
; Sequence 32, Application US/10367978
; GENERAL INFORMATION:
; APPLICANT: GATELY, DENNIS
; TITLE OF INVENTION: PROSTATE SPECIFIC GENES AND THE USE THEREOF IN DESIGN
; FILE REFERENCE: 037003-0301988
; CURRENT APPLICATION NUMBER: US/10/367,978
; CURRENT FILING DATE: 2003-02-19
; PRIOR APPLICATION NUMBER: 60/357,140
; PRIOR FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: 60/396,082
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 60/386,759
; PRIOR FILING DATE: 2002-06-10
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 32
; LENGTH: 1872
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-367-978-32

Query Match 100.0%; Score 518; DB 6; Length 1872;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIRMAQOVFMDTQCSPKTPNNPDHAQSCOLIIELEPPDEKPNQTKSVSPREIVVSL 60
DB 1355 RIRMAQOVFMDTQCSPKTPNNPDHAQSCOLIIELEPPDEKPNQTKSVSPREIVVSL 1414

QY 61 SHOVLTQNLVDIIEEYVKGSPGEKTIQVPEAKLAGFLRYISMOMLAIVPDLIDSYR 120
DB 1415 SHOVLTQNLVDIIEEYVKGSPGEKTIQVPEAKLAGFLRYISMOMLAIVPDLIDSYR 1474

QY 121 TAREFDTPSGIKCLIKKVSIGGAANLYROSAMSFNIFPAVCAVLTNOETITAEQVK 180
DB 1475 TAREFDTPSGIKCLIKKVSIGGAANLYROSAMSFNIFPAVCAVLTNOETITAEQVK 1534

QY 181 VLEFDDRSSTDSQCCSEDEDIFEEETAOVSPPRGKRRQWRAMPPLISQPPSNADWW 240
DB 1535 VLEFDDRSSTDSQCCSEDEDIFEEETAOVSPPRGKRRQWRAMPPLISQPPSNADWW 1594

QY 241 LVKRLHKLQMLCNNTYQWHLIDENCMSEPPIFKGDFFILPSFOSSESTPSTGFSFGKE 300
DB 1595 LVKRLHKLQMLCNNTYQWHLIDENCMSEPPITKGDFFILPSFOSSESTPSTGFSFGKE 1654

QY 301 TPSEDRSQRREHMGESLSLKAGGGDLLPPSPKVEKKDPSRKKEWENAGNKIYTMAD 360

Db 1655 TPSEDDRSOSRHHMESLSLKAGGDLPLPSPKVKEDPSRKKEMWENAKKITWAD 1714
|||
Qy 361 KTISKLTMYKKRKOQHNLSAPFKVEYKKEGPLGPRGDSPLLQRPQHLMDQOMHS 420
|||
Db 1715 KTISKLTMYKKRKOQHNLSAPFKVEYKKEGPLGPRGDSPLLQRPQHLMDQOMHS 1774
|||
Qy 421 PSAGELLARODKRRPSGSGSSLSVSRDAEQIQATNMVLTVINQIILPDQFTALQ 480
|||
Db 1775 PSAGELLARODKRRPSGSGSSLSVSRDAEQIQATNMVLTVINQIILPDQFTALQ 1834
|||
Qy 481 PAVFPCISOLTCHTVDIRVQAVREMLGRGVRYDIIV 518
|||
Db 1835 PAVFPCISOLTCHTVDIRVQAVREMLGRGVRYDIIV 1872
|||

RESULT 2
US-10-462-850-3170
; Sequence 3170, Application US/10462850
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/10/462,850
; CURRENT FILING DATE: 2003-06-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProtPatent version 1.0
; SEQ ID NO 3170
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in Unknown at 6-15 and may interact with Sequenc
US-10-462-850-3170

Query Match 1.4%; Score 7; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 441 SLSVSV 447
|||
Db 4 SLSVSV 10

RESULT 3
US-10-273-573-8627
; Sequence 8627, Application US/10273573
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: US/10/273,573
; CURRENT FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 8627
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1) --(151)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-273-573-8627

Query Match 1.4%; Score 7; DB 6; Length 151;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 269 EPIPIKG 275
|||
Db 140 EPIPIKG 146

RESULT 4
US-09-291-417D-99
; Sequence 99, Application US/09291417D
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHYTE, DAVID
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
; FILE REFERENCE: 038602/0329
; CURRENT APPLICATION NUMBER: US/09/291,417D
; CURRENT FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 60/081,784
; PRIOR FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 99
; LENGTH: 418
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-291-417D-99

Query Match 1.4%; Score 7; DB 5; Length 418;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 425 PELLRD 431
|||
Db 235 PELLRD 241

RESULT 5
US-60-478-196-3273
; Sequence 3273, Application US/60478196
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Lemieux, Sebastien
; APPLICANT: Hu, Meng
; APPLICANT: Roemer, Jerry
; TITLE OF INVENTION: IDENTIFICATION OF ESSENTIAL GENES OF ASPERGILLUS FUMIGATUS AND ME
; FILE REFERENCE: 10182-026-888
; CURRENT APPLICATION NUMBER: US/60/478,196
; CURRENT FILING DATE: 2003-06-13
; NUMBER OF SEQ ID NOS: 4000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3273
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-60-478-196-3273

Query Match 1.4%; Score 7; DB 7; Length 609;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 499 VROAVRE 505
|||
Db 587 VROAVRE 593

RESULT 6
US-10-273-573-10795
; Sequence 10795, Application US/10273573
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066

```

; CURRENT APPLICATION NUMBER: US/10/273,573
; CURRENT FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 10795
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-273-573-10795

```

```

Query Match      1.4%; Score 7; DB 6; Length 642;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      135 LKRVSGI 141
      |||||
Db      122 LKRVSGI 128

```

```

RESULT 7
US-10-294-433-224
; Sequence 224, Application US/10294433
; GENERAL INFORMATION:

```

```

; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 792C1P4
; CURRENT APPLICATION NUMBER: US/10/294,433
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: PCT/US01/14826
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/989,600
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 09/577,408
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: 10/115,831
; PRIOR FILING DATE: 2002-04-02
; PRIOR APPLICATION NUMBER: 09/677,298
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/695,781
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 10/150,802
; PRIOR FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: 09/715,869
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 10/167,379
; PRIOR FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 09/775,330
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 864
; SOFTWARE: Custom
; SEQ ID NO 224
; LENGTH: 1162
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-294-433-224

```

```

Query Match      1.4%; Score 7; DB 6; Length 1162;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      427 LLRODKR 433
      |||||
Db      941 LLRODKR 947

```

```

RESULT 8
US-10-273-573-10799
; Sequence 10799, Application US/10273573
; GENERAL INFORMATION:

```

```

; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: US/10/273,573
; CURRENT FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 10799
; LENGTH: 1246
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (73)..(84)
; OTHER INFORMATION: PROTEIN REPEAT MUSCLE CALCIUM-BI domain identified by
; OTHER INFORMATION: EMATRIX, accession number PD00301A, p-value=6.400e-09, raw score
; OTHER INFORMATION: 10.24
US-10-273-573-10799

```

```

Query Match      1.4%; Score 7; DB 6; Length 1246;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      135 LKRVSGI 141
      |||||
Db      708 LKRVSGI 714

```

```

RESULT 9
US-10-462-850-104
; Sequence 104, Application US/10462850
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/10/462,850
; CURRENT FILING DATE: 2003-06-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProPatent version 1.0
; SEQ ID NO 104
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in BTF3L3 at 25-34 and may interact with Sequence
; OTHER INFORMATION: in this patent.
US-10-462-850-104

```

```

Query Match      1.2%; Score 6; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      434 PRSGST 439
      |||||
Db      1 PRSGST 6

```

```

RESULT 10
US-10-462-850-3172
; Sequence 3172, Application US/10462850
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/10/462,850
; CURRENT FILING DATE: 2003-06-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProPatent version 1.0
; SEQ ID NO 3172

```

```

; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in unknown at 5-14 and may interact with Sequenc
US-10-462-850-3172

Query Match
Best Local Similarity 1.2%; Score 6; DB 6; Length 10;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 441 SLSVS 446
DB 5 SLSVS 10

RESULT 11
US-10-462-850-3684
; Sequence 3684, Application US/10462850
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/10/462,850
; CURRENT FILING DATE: 2003-06-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProtPatent version 1.0
; SEQ ID NO 3684
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in ATOH1 OR ATH1 at 315-324 and may interact with
US-10-462-850-3684

Query Match
Best Local Similarity 1.2%; Score 6; DB 6; Length 10;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 290 TPSTG 295
DB 2 TPSTG 7

RESULT 12
US-10-462-850-3685
; Sequence 3685, Application US/10462850
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/10/462,850
; CURRENT FILING DATE: 2003-06-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProtPatent version 1.0
; SEQ ID NO 3685
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in ATOH1 OR ATH1 at 315-324 and may interact with
US-10-462-850-3685

Query Match
Best Local Similarity 1.2%; Score 6; DB 6; Length 10;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 290 TPSTG 295
DB 2 TPSTG 7
```

```

; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in unknown at 5-14 and may interact with Sequenc
US-10-462-850-3172

Query Match
Best Local Similarity 1.2%; Score 6; DB 6; Length 10;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 441 SLSVS 446
DB 5 SLSVS 10

RESULT 11
US-10-462-850-3684
; Sequence 3684, Application US/10462850
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/10/462,850
; CURRENT FILING DATE: 2003-06-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProtPatent version 1.0
; SEQ ID NO 3684
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in ATOH1 OR ATH1 at 315-324 and may interact with
US-10-462-850-3684

Query Match
Best Local Similarity 1.2%; Score 6; DB 6; Length 10;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 290 TPSTG 295
DB 2 TPSTG 7

RESULT 12
US-10-462-850-3685
; Sequence 3685, Application US/10462850
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/10/462,850
; CURRENT FILING DATE: 2003-06-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProtPatent version 1.0
; SEQ ID NO 3685
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in ATOH1 OR ATH1 at 315-324 and may interact with
US-10-462-850-3685

Query Match
Best Local Similarity 1.2%; Score 6; DB 6; Length 10;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 290 TPSTG 295
DB 2 TPSTG 7

RESULT 13
US-10-369-060A-111
; Sequence 111, Application US/10369060A
; GENERAL INFORMATION:
; APPLICANT: Saliberg, Matti
; TITLE OF INVENTION: SYNTHETIC PEPTIDES THAT BIND TO THE
; FILE REFERENCE: TRIPED.020CPIC1
; CURRENT APPLICATION NUMBER: US/10/369,060A
; CURRENT FILING DATE: 2003-02-14
; PRIOR APPLICATION NUMBER: 09/839,447
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 09/556,605
; PRIOR FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 111
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mus musculus
US-10-369-060A-111

Query Match
Best Local Similarity 1.2%; Score 6; DB 6; Length 23;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 307 RSQSR 312
DB 15 RSQSR 20

RESULT 14
US-10-273-573-7491
; Sequence 7491, Application US/10273573
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: US/10/273,573
; CURRENT FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10394
; SOFTWARE: Custom
; SEQ ID NO 7491
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-273-573-7491

Query Match
Best Local Similarity 1.2%; Score 6; DB 6; Length 40;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 VKGSP 83
DB 16 VKGSP 21

RESULT 15
US-10-273-573-10762
; Sequence 10762, Application US/10273573
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: US/10/273,573
```

```

; CURRENT FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 10762
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (8)..(25)
; OTHER INFORMATION: ISLET AMYLOID PROTEIN (AMYLIN) SIGNATURE domain identified by
; OTHER INFORMATION: EWATRIX, accession number PR00818C, p-value=9.024e-09, raw score
; OTHER INFORMATION: of 6.95
US-10-273-573-10762

```

```

Query Match          1.2%; Score 6; DB 6; Length 97;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      63 QVLLQN 68
        |||||
Db      60 QVLLQN 65

```

Search completed: July 24, 2003, 12:46:18
 Job time : 15.3684 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 12:15:23 ; Search time 34.1288 Seconds

(without alignments)
713.761 Million cell updates/sec

Title: US-09-991-681-27

Perfect score: 518
Sequence: 1 RIRAMAQQVEMLDTCSPKT.....VROAVREMLGRVGRVYDIIV 518

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size : 0

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	1.5	482	1 ODB2_HUMAN	P11182 homo sapien
2	8	1.5	578	1 VATA_METBA	P22662 methanosarc
3	8	1.5	755	1 BPHY_DEIRA	Q9264 deinoxoccus
4	7	1.4	108	1 KVIH_HUMAN	P01605 homo sapien
5	7	1.4	136	1 YBGC_HAEIN	P4679 haemophilus
6	7	1.4	147	1 PR4A_TOBAC	P29062 nicotiana t
7	7	1.4	147	1 PR4B_TOBAC	P29063 nicotiana t
8	7	1.4	158	1 YV45_PASNU	P57900 pasteurella
9	7	1.4	164	1 BUTK_SALTY	Q92408 salmoneilla
10	7	1.4	166	1 OV22_ONCVO	P29779 onchocerca
11	7	1.4	174	1 OV21_ONCVO	P29778 onchocerca
12	7	1.4	182	1 RR4_BELCH	O19990 belamcanda
13	7	1.4	182	1 RR4_HEAMA	O20231 haemaphys
14	7	1.4	182	1 RR4_HYMLI	O20230 itys lutesc
15	7	1.4	182	1 RR4_IRILU	O20235 itys lutesc
16	7	1.4	182	1 RR4_ISOTA	O20236 isophysis t
17	7	1.4	182	1 RR4_LIBRO	O20254 libertia t
18	7	1.4	182	1 RR4_NEOSP	O20251 neomarcia s
19	7	1.4	182	1 RR4_NERBO	O20265 nerine bowd
20	7	1.4	183	1 RR4_ARICA	O19815 aristea cap
21	7	1.4	183	1 RR4_BABST	O19997 babiana str
22	7	1.4	183	1 RR4_CROWN	O20100 crocus nudl
23	7	1.4	183	1 RR4_CROSP	O20103 crocosmia s
24	7	1.4	183	1 RR4_PRESF	O20215 freezia sp.
25	7	1.4	183	1 RR4_LAPNE	O20255 lapeirousia
26	7	1.4	183	1 RR4_SPASP	O20254 spataxis sp
27	7	1.4	194	1 RR4_BOBGL	O19992 bobartia gl
28	7	1.4	194	1 RR4_CYPSP	O20106 cypella sp.
29	7	1.4	194	1 RR4_FURGI	O20213 furcia gi
30	7	1.4	194	1 RR4_IRIEN	O20213 iris enata
31	7	1.4	194	1 RR4_MORST	O20264 moraea spat
32	7	1.4	194	1 RR4_SISST	O20284 sisyrinchiu
33	7	1.4	194	1 RR4_TRISA	O20291 trimezia st

34	7	1.4	195	1 RR4_GLAMU	O20219 gladiolus m
35	7	1.4	195	1 RR4_PILRE	O20278 pillansia t
36	7	1.4	195	1 RR4_ROMRE	O20280 romulea rev
37	7	1.4	195	1 RR4_WATAN	O20399 watsoria an
38	7	1.4	196	1 ECB_ARATH	O07488 arabidopsis
39	7	1.4	196	1 RR4_AGABR	O47025 agave bract
40	7	1.4	196	1 RR4_ASPEB	O47026 asparagus s
41	7	1.4	196	1 RR4_IRIPA	P36461 iris pallid
42	7	1.4	196	1 RR4_NAROD	O47029 narcissus p
43	7	1.4	197	1 RR4_GLAPA	O20221 gladiolus p
44	7	1.4	200	1 CARKB_COTUA	Q98855 coturnix co
45	7	1.4	214	1 BT33_HUMAN	Q13892 homo sapien

ALIGNMENTS

RESULT 1
ODB2_HUMAN STANDARD; PRT; 482 AA.
ID ODB2_HUMAN
AC P11182;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Liponamide acyltransferase component of branched-chain alpha-keto acid dehydrogenase complex, mitochondrial precursor (EC 2.3.1.-) (E2)
DE (b)hydrolypoamide branched chain transacylase (BCKAD E2 subunit).
GN DPT OR BCAT2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=93041936; PubMed=1420314;
RA Lau K.S., Chuang J.L., Herring W.J., Danner D.J., Cox R.P.,
RA Chuang D.T.;
RT "The complete cDNA sequence for dihydrolypoyl transacylase (E2) of human branched-chain alpha-keto acid dehydrogenase complex.";
RL Biochim. Biophys. Acta 1132:319-321(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88198156; PubMed=3245861;
RA Hummel K.B., Litwer S., Bradford A.P., Aitken A., Danner D.J.,
RA Yeaman S.J.;
RT "Nucleotide sequence of a cDNA for branched chain acyltransferase with analysis of the deduced protein structure.";
RL J. Biol. Chem. 263:6165-6168(1988).
RN [3]
RP REVISIONS.
RX MEDLINE=89214230; PubMed=2708389;
RA Danner D.J., Litwer S., Herring W.J., Pruckler J.;
RT "Construction and nucleotide sequence of a cDNA encoding the full-length preprotein for human branched chain acyltransferase.";
RL J. Biol. Chem. 264:7742-7746(1989).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=89302075; PubMed=2742576;
RA Nobukuni Y., Mitsubuchi H., Endo F., Matsuda I.;
RT "Complete primary structure of the transacylase (E2b) subunit of the human branched chain alpha-keto acid dehydrogenase complex.";
RL Biochem. Biophys. Res. Commun. 161:1035-1041(1989).
RN [5]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Siemen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.T., Wang J., Heien F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullan S.J.,
RA Bosak S.A., McEwen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hilly S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W.,
RA Buterfield Y.S.N., Krzywinski M.I., Skalska M., Smalins D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[6]
RN SEQUENCE OF 1-313 FROM N.A.
RX MEDLINE=88241022; PubMed=2837277;
RA Lau K.S., Griffin T.A., Hu C.-W.C., Chuang D.T.;
RT "Conservation of primary structure in the lipoyl-bearing and
RT dihydrolioyl dehydrogenase binding domains of mammalian
RT branched-chain alpha-keto acid dehydrogenase complex: molecular
RT cloning of human and bovine transacylase (E2) cDNAs.";
RL Biochemistry 27:1972-1981(1988).
[7]
RN VARIANT MSUD CVS-276.
RX MEDLINE=91128420; PubMed=1847055;
RA Fisher C.W., Lau K.S., Fisher C.R., Wynn R.M., Cox R.P., Chuang D.T.;
RT "A 17-bp insertion and a Phe215-->Cys missense mutation in the
RT dihydrolioyl transacylase (E2) mRNA from a thiamine-responsive maple
RT syrup urine disease patient WG-34.";
RL Biochem. Biophys. Res. Commun. 174:804-809(1991).
RN [8]
RN VARIANTS MSUD MET-98 AND SER-384.
RX MEDLINE=96284532; PubMed=9621512;
RA Tsunuma M., Mitsubuchi H., Maruy S., Miura Y., Hayashida Y.,
RA Kinugasa A., Ishitsu T., Matsuda I., Indo Y.;
RT "Molecular basis of intermittent maple syrup urine disease: novel
RT mutations in the E2 gene of the branched-chain alpha-keto acid
RT dehydrogenase complex.";
RL Hum. Genet. 43:91-100(1998).
CC -1- FUNCTION: THE BRANCHED-CHAIN ALPHA-KETO DEHYDROGENASE COMPLEX
CC CATALYZES THE OVERALL CONVERSION OF ALPHA-KETO ACIDS TO ACYL-COA
CC AND CO(2). IT CONTAINS MULTIPLE COPIES OF 3 ENZYMAIC COMPONENTS:
CC BRANCHED-CHAIN ALPHA-KETO ACID DECARBOXYLASE (E1), LIPOMIDE
CC ACYLTRANSFERASE (E2) AND LIPOMIDE DEHYDROGENASE (E3).
CC -1- COFACTOR: THE E2 COMPONENT CONTAINS ONE COVALENTLY-BOUND LIPOYL
CC COFACTOR.
CC -1- SUBUNIT: FORMS A 24-POLYPEPTIDE STRUCTURAL CORE WITH OCTAHEDRAL
CC SYMMETRY.
CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -1- DISEASE: Defects in DBT are a cause of maple syrup urine disease
CC (MSUD); an autosomal recessive disorder characterized by mental
CC and physical retardation, feeding problems, and a maple syrup odor
CC to the urine.
CC -1- MISCELLANEOUS: THE CATALYTIC FUNCTION OF THIS ENZYME IS TO ACCEPT,
CC AND TO TRANSFER TO COENZYME A, ACYL GROUPS THAT ARE GENERATED BY
CC THE BRANCHED-CHAIN ALPHA-KETO ACID DECARBOXYLASE COMPONENT.
CC -1- SIMILARITY: BELONGS TO THE 2-OXOACID DEHYDROGENASE FAMILY.
CC -1- SIMILARITY: Contains 1 lipoyl-binding domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, X66785; CAA47285.1; -;
CC EMBL, J03208; AAA35589.1; ALT INIT.
CC EMBL, M27093; AAA64512.1; ALT INIT.
CC EMBL, BC016675; AAA16675.1; -;
CC EMBL, M19301; AAA59200.1; ALT_SEQ.

DR PIR; A32422; A32422.
DR PDB; 1K8W; 04-DEC-02.
DR PDB; 1K8O; 14-NOV-01.
DR Genew; HGNC:6998; DBT.
DR GK; P1182; -;
DR MIM; 248410; -;
DR GO; GO:0005947; C:alpha-ketoglutarate dehydrogenase complex (. . .); TAS.
DR GO; GO:0005739; C:mitochondrion; TAS.
DR InterPro; IPR001078; 2-oxoacid dh.
DR InterPro; IPR000089; Biotin lipoyl.
DR InterPro; IPR004167; E3 binding.
DR InterPro; IPR003016; lipoyl.
DR Pfam; PF00198; 2-oxoacid dh; 1.
DR Pfam; PF00364; biotin_lipoyl; 1.
DR Pfam; PF02817; e3_binding; 1.
DR Pfam; PF00115; 2-oxoacid dh; 1.
DR PROSITE; PS00189; LIPOYL; 1.
KW Transferrase; Acyltransferase; Mitochondrion; Transit peptide; Lipoyl;
KW Disease mutation; Maple syrup urine disease; 3D-structure.
FT TRANSIT 1 61
FT CHAIN 62 482
FT FT
FT BINDING 105 105
FT ACT_SITE 452 452
FT ACT_SITE 456 456
FT VARIANT 98 98
FT VARIANT 276 276
FT VARIANT 384 384
FT VARIANT 321 321
FT CONFLICT 354 354
SQ SEQUENCE 482 AA; 53487 MW; A7CA728C8F3D126 CRC64;
Query Match 1.5%; Score 8; DB 1; Length 482;
Best Local Similarity 100.0%; Pred. No. 8.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 329 LPSPKYE 336
Db 217 LPSPKYE 224
RESULT 2
VATA METBA STANDARD; PRT; 578 AA.
AC P22662;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE V-type ATP synthase alpha chain (EC 3.6.3.14) (V-type ATPase subunit
DE A).
GN ATPA.
OS Methanosaeta barkeri.
OC Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;
OC Methanosarcinales; Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2208;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89291829; PubMed=2544575;
RA Inatomi K.I., Eya S., Maeda M., Futai M.;
RT "Amino acid sequence of the alpha and beta subunits of Methanosarcina
RT barkeri ATPase deduced from cloned genes. Similarity to subunits of
RT eukaryotic vacuolar and F0F1-ATPases.";
RL J. Biol. Chem. 264:10954-10959(1989).
CC -1- FUNCTION: PRODUCES ATP FROM ADP IN THE PRESENCE OF A PROTON
CC GRADIENT ACROSS THE MEMBRANE. THE ARCHAEL ALPHA CHAIN IS A
CC CATALYTIC SUBUNIT.
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (in) = ADP + phosphate +
CC H(+) (out).
CC -1- SIMILARITY: Belongs to the ATPase alpha/beta chains family.

```

CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on the way
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC
DR EMBL, J04836; AAA72215.1; -.
DR PIR, A34283; A34283.
DR HAMAP, MF_00309; -, 1.
DR InterPro: IPR005726; A_ATPase_A.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR000793; ATPase_a/b/c.
DR InterPro: IPR000194; ATPase_a/b/centr.
DR InterPro: IPR004100; ATPase_a/b/N.
DR Pfam: PF00006; ATP-synt_ab; 1.
DR Pfam: PF00306; ATP-synt_ab_C; 1.
DR Pfam: PF02874; ATP-synt_ab_N; 1.
DR SMART, SM00382; AAA; 1.
DR TIGRFAMs, TIGR01043; ATP_syn_A_arch; 1.
DR PROSITE, PS00152; ATPase_ALPHA_ARCH; 1.
KM Hydrolyase; ATP synthesis; Hydrogen ion transport; ATP-binding.
FT NP_BIND 228 235 ATP (BY SIMILARITY).
SQ SEQUENCE 578 AA; 63638 MW; 195019833C89552B CRC64;

OY 438 STGSSLSLV 445
Db 66 STGSSLSLV 73

RESULT 3
BPHY_DEIRA
ID_BPHY_DEIRA STANDARD; PRT; 755 AA.
AC Q9RZA4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Bacteriophytochrome (EC 2.7.3.-) (Phytochrome-like protein).
GN BpHP OR DRA0050.
OS Deinococcus radiodurans.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcaceae.
OC NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1 / ATCC 13939 / DSM 20539 / NCIB 9279;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1."
RL Science 286:1571-1577 (1999).
RN [2]
RP CHARACTERIZATION, AND MUTAGENESIS OF MET-259, CYS-289 AND HIS-260.
RX MEDLINE=20085458; PubMed=10617469;
RA Davis S.J., Vener A.V., Vierstra R.D.;
RT "Bacteriophytochromes: phytochrome-like photoreceptors from
RT nonphotosynthetic eubacteria."
RL Science 286:2517-2520 (1999).

-1- FUNCTION: PHOTORECEPTOR WHICH EXISTS IN TWO FORMS THAT ARE
REVERSIBLY INTERCONVERTIBLE BY LIGHT. THE R FORM THAT ABSORBS
MAXIMALLY IN THE RED REGION OF THE SPECTRUM AND THE FR FORM THAT

```

```

CC ASSORES MAXIMALLY IN THE FAR-RED REGION. HAS ALSO A SLIGHT BLUE
CC SHIFT FOR THE FAR-RED MAXIMUM. COULD ALSO ABSORB GREEN LIGHT. MAY
CC PARTICIPATE IN REGULATING PIGMENT SYNTHESIS LIKE THE CAROTENOID
CC DEINOXANTHIN WHICH COULD PROTECT THE BACTERIUM FROM INTENSE
CC VISIBLE LIGHT.
CC
CC -1- PFM: CONTAINS ONE COVALENTLY LINKED TETRAPYRROLE CHROMOPHORE.
CC LACKS THE CONSERVED CYSTEINE WHICH BINDS CHROMOPHORE IN HIGHER
CC PLANTS PHYTOCHROMES; THE ADJACENT HISTIDINE RESIDUE PROBABLY SERVE
CC AS A LIGATION SITE. ABLE TO BIND PHYCOERYTHROBLIN (PEB) AND
CC PHYTOCHROMOBILIN IN VITRO, BUT THE IDENTITY OF THE NATURAL
CC CHROMOPHORE IS UNKNOWN.
CC
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE PHYTOCHROME
CC FAMILY.
CC
CC -1- SIMILARITY: Contains 1 PAS (PER-ARNT-SIM) dimerization domain.
CC
CC -1- SIMILARITY: Contains 1 histidine kinase domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
CC
CC EMBL; AE001862; AAF1261.1; -.
CC PIR; D75598; D75598.
CC TIGR; DRA0050; -.
CC
CC InterPro: IPR003594; AtPbind ATPase.
CC InterPro: IPR004358; Bact_sens_pr_C.
CC InterPro: IPR003018; GAF.
CC InterPro: IPR003661; His_kinA.
CC InterPro: IPR005467; His_kinase.
CC InterPro: IPR000014; PAS domain.
CC InterPro: IPR001294; Phytochrome.
CC
CC Pfam; PF01590; GAF.1
CC Pfam; PF02518; HATPase_C.1.
CC Pfam; PF00512; HisKA.1.
CC Pfam; PF00360; Phytochrome.1.
CC PRINTS; PR00344; BCTRLSENSOR.
CC PRINTS; PR01033; PHYTOCHROME.
CC SMART; SMO0065; GAF.1.
CC SMART; SMO0387; HATPase_C.1.
CC SMART; SMO0388; HisKA.1.
CC SMART; SMO0091; PAS.1.
CC
CC DR PROSITE; PS50109; HIS_KIN.1.
CC DR PROSITE; PS50112; PAS; FALSE_NEG.
CC DR PROSITE; PS50245; PHYTOCHROME.1; FALSE_NEG.
CC DR PROSITE; PS50046; PHYTOCHROME.2; 1.
CC
CC KM Sensory transduction; Transferrase, kinase; Phosphorylation;
CC Photoreceptor; Phytochrome; Chromophore; Complete proteome.
CC
CC FT DOMAIN 26 94
CC FT DOMAIN 95 504
CC FT DOMAIN 529 747
CC FT BINDING 260 260
CC FT MOD_RES 532 532
CC FT MUTAGEN 259 259
CC
CC FT MUTAGEN 259 259
CC
CC FT MUTAGEN 260 260
CC
CC FT MUTAGEN 289 289
CC
CC FT SEQUENCE 755 AA; 81584 MW; A631E471B208F187 C6C64;
CC
CC Query Match 1.5%; Score 8; DB 1; Length 755;
CC Best Local Similarity 100.0%; Pred.No.13;
CC Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

RESULT 4
KVIM_HUMAN STANDARD; PRT; 108 AA.
AC P01605;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG kappa chain V-I region lay.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=77038198; PubMed=824717;
RA Capra J.D., Klapper D.G.;
RT "Complete amino acid sequence of the variable domains of two human
RT IGM anti-gamma globulins (Iay/Pom) with shared idiotypic
RT specificities."
RL Stand. J. Immunol. 5:677-684(1976).
CC -I- MISCELLANEOUS: THE SECOND AND THIRD HYPERVARIABLE REGIONS OF THIS
CC CHAIN ARE IDENTICAL WITH THOSE OF THE HUMAN POM V-III KAPPA CHAIN,
CC WITH WHICH IT SHARES CERTAIN IDIOTYPIC DETERMINANTS.
CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGM WITH ANTI-GAMMA
CC GLOBULIN ACTIVITY.
DR PIR: A01871; KIHILY.
DR HSSP: P01607; IREI.
DR GO: GO:0005576; C:extracellular; NAS.
DR GO: GO:0003823; F:antigen binding activity; NAS.
DR GO: GO:0006955; P:immune response; NAS.
DR InterPro: IPR003006; Ig-like.
DR InterPro: IPR003596; Ig_v.
DR Pfam: PF00047; Ig_1.
DR SMART: SM00406; IGV; 1.
DR PROSITE: PS00835; IG LIKE; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 23 FRAMEWORK-1.
FT DOMAIN 2 34 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 3 49 FRAMEWORK-2.
FT DOMAIN 4 56 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 5 57 FRAMEWORK-3.
FT DOMAIN 6 88 FRAMEWORK-3.
FT DOMAIN 7 89 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 8 97 FRAMEWORK-4.
FT DISULFID 23 88 BY SIMILARITY.
FT NON_TER 108
SQ SEQUENCE 108 AA; 11834 MW; 73993A95431434A CRC64;

Query Match 1.4%; Score 7; DB 1; Length 108;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 441 SLSVSV 447
DB 9 SLSVSV 15

RESULT 5
YBGC_HAEIN STANDARD; PRT; 136 AA.
AC P44679;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein HI0386.
GN HI0386.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shiley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Frazer C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd."
RL Science 269:496-512(1995).
CC -I- SIMILARITY: BELONGS TO THE 4-HYDROXYBENZOYL-COA THIOESTERASE
CC FAMILY. STRONG, TO E.COLI YBGC.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use. By non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL: U32722; AAC2204.1; -.
DR PIR: H64150; H64150.
DR TIGR: H10386; -.
DR InterPro: IPR006684; 4HBCOA_thioesterase.
DR InterPro: IPR006683; Thioestr_sufp.
DR Pfam: PF03061; 4HBT; 1.
DR TIGRFams: TIGR00051; TIGR00051; 1.
DR PROSITE: PS01328; 4HBCOA_THIOESTERASE; 1.
KW Hypothetical protein; Hydrolyase; Complete proteome.
FT ACT_SITE 18 18 BY SIMILARITY.
SQ SEQUENCE 136 AA; 15597 MW; E1D5D2E3F800B261 CRC64;

Query Match 1.4%; Score 7; DB 1; Length 136;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 381 AFPEVK 387
DB 123 AFPEVK 129

RESULT 6
PRA4_TOBAC STANDARD; PRT; 147 AA.
AC P29052;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Pathogenesis-related protein PR-4a precursor.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamiales; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Xanthi; TISSUE=leaf;
RX MEDLINE=92079884; PubMed=174523;
RA Friedrich L., Moyer W., Ward E., Ryals J.;
RT "Pathogenesis-related protein 4 is structurally homologous to the
RT carboxy-terminal domains of hevein, win-1 and win-2."
RL Mol. Gen. Genet. 230:113-119(1991).
CC -I- INDUCTION: By TMV infection.
CC -I- SIMILARITY: BELONGS TO THE BARWIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X60281; CAA42820.1; -.
DR PIR; S18598; S18598.
DR HSSP; P28814; 1B4.
DR InterPro; IPR001153; Barwin.
DR Pfam; PF00967; Barwin.1.
DR PRINTS; PR00602; BARWIN.1.
DR ProDom; PD004535; Barwin.1.
DR PROSITE; PS00771; BARWIN_1; 1.
DR PROSITE; PS00772; BARWIN_2; 1.
KM Plant defense; Pathogenesis-related protein; Signal.
FT SIGNAL 1 25
FT CHAIN 26 147
FT DISULFID 54 86
FT DISULFID 75 109
FT DISULFID 89 145
SQ SEQUENCE 147 AA; 16221 MW; 80060A2AFB40D30 CRC64;

Query Match 1.4%; Score 7; DB 1; Length 147;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 396 GPRGQDS 402
Db 79 GPRGQDS 85

RESULT 7
PRAB_TOBAC STANDARD; PRT; 147 AA.
ID PRAB_TOBAC STANDARD; PRT; 147 AA.
AC P29065;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Pathogenesis-related protein PR-4B precursor.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Assteridae; Lamiales; Solanales; Solanaceae; Nicotiana.
OC NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Xanthi; TISSUE=leaf;
RX MEDLINE=92079884; PubMed=1745223;
RA Friedrich L., Moyer M., Ward E., Ryals J.;
RT "Pathogenesis-related protein 4 is structurally homologous to the
RT carboxy-terminal domains of hevein, Win-1 and Win-2.";
RL Mol. Genet. 230:113-119(1991).
CC -1- INDUCTION: By TMV infection.
CC -1- SIMILARITY: BELONGS TO THE BARWIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X60282; CAA42821.1; -.
DR HSSP; P28814; 1B4.
DR InterPro; IPR001153; Barwin.
DR Pfam; PF00967; Barwin.1.
DR PRINTS; PR00602; BARWIN.1.
DR ProDom; PD004535; Barwin.1.
DR PROSITE; PS00771; BARWIN_1; 1.
DR PROSITE; PS00772; BARWIN_2; 1.
KM Plant defense; Pathogenesis-related protein; signal.
FT SIGNAL 1 25
FT SIGNAL 25

```

```

FT CHAIN 26 147
FT DISULFID 54 86
FT DISULFID 75 109
FT DISULFID 89 145
SQ SEQUENCE 147 AA; 16235 MW; FE7579E1FB2C874 CRC64;

Query Match 1.4%; Score 7; DB 1; Length 147;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 396 GPRGQDS 402
Db 79 GPRGQDS 85

RESULT 8
YA45_PASMU STANDARD; PRT; 158 AA.
ID YA45_PASMU STANDARD; PRT; 158 AA.
AC P57900;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical UPF0054 protein PM1045.
GN PM1045.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OC NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pm70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whitram T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -1- SIMILARITY: Belongs to the UPF0054 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE006145; AA03129.1; -.
DR HAMAP; MF_00009; -; 1.
DR InterPro; IPR002036; UPF0054.
DR Pfam; PF02130; UPF0054; 1.
DR ProDom; PD005688; UPF0054; 1.
DR TIGRFAMs; TIGR00043; TIGR00043; 1.
DR PROSITE; PS01306; UPF0054; 1.
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 158 AA; 17951 MW; FFB6D9B63699250A CRC64;

Query Match 1.4%; Score 7; DB 1; Length 158;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 451 EAQIQAW 457
Db 21 EAQIQAW 27

RESULT 9
EUTK_SALTY STANDARD; PRT; 164 AA.
ID EUTK_SALTY STANDARD; PRT; 164 AA.
AC Q92F08;
DT 30-MAY-2000 (Rel. 39, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ethanolamine utilization protein eutk precursor.
GN EUTK OR STM2455.

```

```

OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
CC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2;
RX MEDLINE=99395039; PubMed=10464203;
RA Kofoed E.C., Rappleye C.A., Stojiljkovic I., Roth J.R.;
RT "The 17-gene ethanolamine (eut) operon of Salmonella typhimurium
encodes five homologues of carboxysome shell proteins.";
RL Bacteriol. 181:5317-5329(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
Ryan E., Sun H., Florea L., Miller W., Stuckey T., Nhan M.,
Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
LT2.";
RL Nature 413:852-856(2001).
CC -1- FUNCTION: MAY BE INVOLVED IN THE FORMATION OF A SPECIFIC
MICROCOMPARTMENT IN THE CELL IN WHICH THE METABOLISM OF
POTENTIALITY TOXIC BY-PRODUCTS TAKES PLACE.
CC -1- PATHWAY: Ethanolamine utilization.
CC -1- SIMILARITY: BELONGS TO THE CCMR/CCMR/COSI/PDUA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF093749; AAC78126.1; -.
CC EMBL; AE008810; AAL21349.1; -.
CC PDB; 1LGO; 01-MAY-02.
CC StyGene; SG10630; outK.
CC InterPro; IPR000249; Bact_microcomp.
CC Pfam; PF00936; Bact_microcomp.
CC ProDom; PD003442; Bact_microcomp; 1.
CC PROSITE; PS01139; BACT_MICROCOMP; 1.
CC KMW Signal; Complete proteome; 3D-structure.
CC FT SIGNAL 1 16 POTENTIAL.
CC CHAIN 17 164 ETHANOLAMINE UTILIZATION PROTEIN EUTK.
CC CONFLICT 65 65 O -> K (IN REF. 1).
CC CONFLICT 153 154 LR -> FG (IN REF. 1).
CC SEQUENCE 164 AA; 17486 MW; 9CA9F2DA4EACF2BC CRC64;

Query Match 1.4%; Score 7; DB 1; Length 164;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 59 LLSHVL 65
DB 31 LLSHVL 37

RESULT 10
OV22_ONCVO STANDARD; PRT; 166 AA.
AC P29779;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE Small heat shock protein OV25-2 (Fragment).
GN OV25-2.
OS Onchocerca volvulus.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;

```

```

CC Onchocercidae; Onchocerca.
OX NCBI_TaxID=6282;
RN [1]
RP SEQUENCE FROM N.A.
RC Hoeffle W.;
RT Submitted (OCT-1992) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SMALL HEAT SHOCK PROTEIN (HSP20)
FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X68669; CAA4863.1; -.
CC InterPro; IPR002068; Hsp20.
CC Pfam; PF00011; HSP20; 1.
CC PROSITE; PS01031; HSP20; 1.
CC KMW Heat shock.
CC FT NON TER 1
CC SEQUENCE 166 AA; 18636 MW; 658690B4D0B946CD CRC64;

Query Match 1.4%; Score 7; DB 1; Length 166;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 443 LSVSVRD 449
DB 71 LSVSVRD 77

RESULT 11
OV21_ONCVO STANDARD; PRT; 174 AA.
AC P29778;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE Small heat shock protein OV25-1.
GN OV25-1.
OS Onchocerca volvulus.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC Onchocercidae; Onchocerca.
OX NCBI_TaxID=6282;
RN [1]
RP SEQUENCE FROM N.A.
RC Hoeffle W.;
RT Submitted (OCT-1992) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SMALL HEAT SHOCK PROTEIN (HSP20)
FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X68668; CAA4863.1; -.
CC InterPro; IPR001436; Crystallin_alpha.
CC InterPro; IPR002068; Hsp20.
CC Pfam; PF00011; HSP20; 1.
CC PRINTS; PR00299; ACRYSTALLIN.
CC PROSITE; PS01031; HSP20; 1.
CC KMW Heat shock.
CC SEQUENCE 174 AA; 19935 MW; 81F8BA4ADAC6F5FE CRC64;

Query Match 1.4%; Score 7; DB 1; Length 174;
Best Local Similarity 100.0%; Pred. No. 35;

```

Matches	7;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Oy	443	LSVSVRD	449						
Db	83	LSVSVRD	89						
RESULT 12									
IR4_BELCH		STANDARD;		PRT;	182	AA.			
AC	019990;								
DT	15-DEC-1998	(Rel. 37, Created)							
DT	15-DEC-1998	(Rel. 37, Last sequence update)							
DT	28-FEB-2003	(Rel. 41, Last annotation update)							
DE	Chloroplast	30S ribosomal protein S4 (Fragment).							
GN	RP94.								
OS	Belamcanda chinensis.								
OC	Chloroplast.								
OC	Eukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;								
OC	Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Iridaceae;								
OC	Belamcanda.								
OX	NCBI_TaxID=58944;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RA	Souza-Chies T.T., Bittar G., Nadot S., Carter L., Besin E.,								
RA	Lejeune B.P.;								
RT	"Phylogenetic analysis of Iridaceae with parsimony and distance								
RT	methods using the plastid gene rps4."								
RT	Plant Syst. Evol. 204:109-123(1997).								
CC	-1- FUNCTION: One of the primary rRNA binding proteins. It binds								
CC	directly to 16S rRNA where it nucleates assembly of the body of								
CC	the 30S subunit (By similarity).								
CC	-1- FUNCTION: With S5 and S12 plays an important role in translational								
CC	accuracy (By similarity).								
CC	-1- SUBUNIT: Part of the 30S ribosomal subunit. Contacts protein S5.								
CC	The interaction surface between S4 and S5 is involved in control								
CC	of translational fidelity (By similarity).								
CC	-1- SUBCELLULAR LOCATION: Chloroplast.								
CC	-1- SIMILARITY: Contains 1 S4 RNA-binding domain.								
CC	-1- SIMILARITY: BELONGS TO THE S4P FAMILY OF RIBOSOMAL PROTEINS.								
CC	-----								
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration								
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -								
CC	the European Bioinformatics Institute. There are no restrictions on								
CC	use by non-profit institutions as long as its content is in no way								
CC	modified and this statement is not removed. Usage by and for commercial								
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/								
CC	or send an email to license@isb-sib.ch).								
CC	-----								
DR	EMBL; 268235; CAA92533.1; -.								
DR	HSSP; P81288; IC05.								
DR	HAMAP; MF_01306; -; 1.								
DR	InterPro; IPR001912; Ribosomal_S4.								
DR	InterPro; IPR002942; S4.								
DR	InterPro; IPR005709; S4_bact_ORG.								
DR	Pfam; PF00163; Ribosomal_S4; 1.								
DR	Pfam; PFO1479; S4; 1.								
DR	SMART; SMO0363; S4; 1.								
DR	TIGRFAMs; TIGR01017; rpsd_bact; 1.								
DR	PROSITE; PS00632; RIBOSOMAL_S4; 1.								
DR	PROSITE; PS50889; S4; 1.								
KW	Ribosomal protein; RNA-binding; rRNA-binding; Chloroplast.								
FT	NON TER	1							
FT	DOMAIN	82	143			S4 RNA-BINDING.			
FT	NON TER	182							
SO	SEQUENCE	182	AA;	21048	MM;	E3620B0641A04EFS	CRC64;		
Query Match 1.4%; Score 7; DB 1; Length 182;									
Best Local Similarity 100.0%; Pred. No. 36;									
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Oy	432	KRPRSGS	438						

```

RESULT 13
ID RR4_HEAMA STANDARD; PRT; 182 AA.
AC 020231;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chloroplast 30S ribosomal protein S4 (Fragment).
GN RPS4.
OS Haemanthus magnificae.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Amaryllidaceae;
OC Haemanthus.
CX NCBI_TaxID=59038;
RN [1]
RP SEQUENCE FROM N.A.
RA Souza-Chies T.T., Bitar G., Nadot S., Carter L., Beein E.,
RA Lejeune B.P.;
RT "Phylogenetic analysis of Iridaceae with parsimony and distance
RT methods using the plastid gene rps4."
RL Plant Syst. Evol. 204:109-123(1997).
CC -1- FUNCTION: One of the primary rRNA binding proteins, it binds
CC directly to 16S rRNA where it nucleates assembly of the body of
CC the 30S subunit (By similarity).
CC -1- FUNCTION: With S5 and S12 plays an important role in translational
CC accuracy (By similarity).
CC -1- SUBUNIT: Part of the 30S ribosomal subunit. Contacts protein S5.
CC The interaction surface between S4 and S5 is involved in control
CC of translational fidelity (By similarity).
CC -1- SUBCELLULAR LOCATION: Chloroplast.
CC -1- SIMILARITY: Contains 1 S4 RNA-binding domain.
CC -1- SIMILARITY: BELONGS TO THE S4P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on ways
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; Z68253; CAA92551.1; -.
DR HSSP; P81288; IC05.
DR HAMAP; MF_01306; -; 1.
DR InterPro; IPR001912; Ribosomal_S4.
DR InterPro; IPR002942; S4.
DR InterPro; IPR005709; S4 bact.org.
DR Pfam; PF00163; Ribosomal_S4_1.
DR Pfam; PF01479; S4; 1.
DR SMART; SM00363; S4; 1.
DR TIGRFAMs; TIGR01017; rpsd_bact; 1.
DR PROSITE; PS00632; RIBOSOMAL_S4; 1.
DR PROSITE; PS50889; S4; 1.
KW Ribosomal protein; rRNA-binding; rRNA-binding; Chloroplast.
FT NON TER 1 1
FT DOMAIN 82 143 S4 RNA-BINDING.
FT NON TER 182 182
SQ SEQUENCE 182 AA; 21086 MW; E091810DFDFED3B CRC64;

Query Match 1.4%; Score 7; DB 1; Length 182;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 432 KRPSGS 438
DB 17 KRPSGS 23

```

```

RR4_HYMLI
ID RR4_HYMLI STANDARD: PRT, 182 AA.
AC 020230;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chloroplast 30S ribosomal protein S4 (Fragment).
GN RPS4.
OS Hymenocallis littoralis (Beach spiderlily).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Amaryllidaceae;
OC Hymenocallis.
OC NCBI_TaxID=59040;
OX NCBI_TaxID=59040;
RN [1]
RP SEQUENCE FROM N.A.
RA Souza-Chies T.T., Bittar G., Nadot S., Carter L., Besin E.,
RA Lejeune B.P.;
RT "Phylogenetic analysis of Iridaceae with parsimony and distance
RT methods using the plastid gene rps4."
RL Plant Syst. Evol. 204:109-123(1997).
CC -1- FUNCTION: One of the primary rRNA binding proteins, it binds
CC directly to 16S rRNA where it nucleates assembly of the body of
CC the 30S subunit (By similarity).
CC -1- FUNCTION: With S5 and S12 plays an important role in translational
CC accuracy (By similarity).
CC -1- SUBUNIT: Part of the 30S ribosomal subunit. Contacts protein S5.
CC The interaction surface between S4 and S5 is involved in control
CC of translational fidelity (By similarity).
CC -1- SUBCELLULAR LOCATION: Chloroplast.
CC -1- SIMILARITY: Contains 1 S4 RNA-binding domain.
CC -1- SIMILARITY: BELONGS TO THE S4P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z68251; CAA92549.1; -.
DR HSSP; P81288; IC05.
DR HAMAP; MF_01306; -.
DR InterPro; IPR001912; Ribosomal_S4.
DR InterPro; IPR002942; S4.
DR Pfam; PF00163; Ribosomal_S4; 1.
DR Pfam; PF01479; S4; 1.
DR SMART; SM00363; S4; 1.
DR TIGRfams; TIGR01017; rpsd_bact; 1.
DR PROSITE; PS00632; RIBOSOMAL_S4; 1.
DR PROSITE; PS50889; S4; 1.
KW Ribosomal protein; RNA-binding; rRNA-binding; Chloroplast.
FT DOMAIN 1 1
FT NON_TER 82 143 S4 RNA-BINDING.
FT NON_TER 182 182
SQ SEQUENCE 182 AA; 21050 MW; 78B1261C3F6DC8AF CRC64;
OY 432 KRPRSGS 438
Db 17 KRPRSGS 23

```

```

DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chloroplast 30S ribosomal protein S4 (Fragment).
GN RPS4.
OS Iris lutescens.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Iridaceae;
OC Iris.
OC NCBI_TaxID=58956;
OX NCBI_TaxID=58956;
RN [1]
RP SEQUENCE FROM N.A.
RA Souza-Chies T.T., Bittar G., Nadot S., Carter L., Besin E.,
RA Lejeune B.P.;
RT "Phylogenetic analysis of Iridaceae with parsimony and distance
RT methods using the plastid gene rps4."
RL Plant Syst. Evol. 204:109-123(1997).
CC -1- FUNCTION: One of the primary rRNA binding proteins, it binds
CC directly to 16S rRNA where it nucleates assembly of the body of
CC the 30S subunit (By similarity).
CC -1- FUNCTION: With S5 and S12 plays an important role in translational
CC accuracy (By similarity).
CC -1- SUBUNIT: Part of the 30S ribosomal subunit. Contacts protein S5.
CC The interaction surface between S4 and S5 is involved in control
CC of translational fidelity (By similarity).
CC -1- SUBCELLULAR LOCATION: Chloroplast.
CC -1- SIMILARITY: Contains 1 S4 RNA-binding domain.
CC -1- SIMILARITY: BELONGS TO THE S4P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z68242; CAA92540.1; -.
DR HSSP; P81288; IC05.
DR HAMAP; MF_01306; -.
DR InterPro; IPR001912; Ribosomal_S4.
DR InterPro; IPR002942; S4.
DR InterPro; IPR005709; S4_bact_org.
DR Pfam; PF00163; Ribosomal_S4; 1.
DR Pfam; PF01479; S4; 1.
DR SMART; SM00363; S4; 1.
DR TIGRfams; TIGR01017; rpsd_bact; 1.
DR PROSITE; PS00632; RIBOSOMAL_S4; 1.
DR PROSITE; PS50889; S4; 1.
KW Ribosomal protein; RNA-binding; rRNA-binding; Chloroplast.
FT DOMAIN 1 1
FT NON_TER 82 143 S4 RNA-BINDING.
FT NON_TER 182 182
SQ SEQUENCE 182 AA; 20996 MW; 193292203A86C6C6 CRC64;
OY 432 KRPRSGS 438
Db 17 KRPRSGS 23

```

Search completed: July 24, 2003, 12:22:04
 Job time : 35.1288 secs

RESULT 15
 RR4_IRITU STANDARD: PRT, 182 AA.
 AC 020235;
 DT 15-DEC-1998 (Rel. 37, Created)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 12:16:59 ; Search time 62.1903 Seconds
(without alignments)
801.015 Million cell updates/sec

Title: US-09-991-681-27

Perfect score: 518
Sequence: 1 RIRAMQOVFMDTQCSPKT.....VRQAVREMLGRVHYDIIV 518

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283308 seqs, 96168682 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	1.7	108	2	Ig kappa chain V-J
2	9	1.7	216	2	competence protein
3	9	1.7	216	2	competence protein
4	8	1.5	427	2	hypothetical prote
5	8	1.5	482	2	dihydroilpoamide S
6	8	1.5	483	2	AD1223
7	8	1.5	578	2	hypothetical prote
8	8	1.5	610	2	H+-transporting tw
9	8	1.5	621	2	hypothetical prote
10	8	1.5	755	2	conserved hypotnet
11	8	1.5	1532	2	photoreceptor - De
12	8	1.4	82	2	hypothetical prote
13	7	1.4	86	2	probable DNA-bindi
14	7	1.4	108	1	hypothetical trans
15	7	1.4	129	2	Ig kappa chain V-I
16	7	1.4	136	2	hypothetical prote
17	7	1.4	137	2	conserved hypotnet
18	7	1.4	139	2	sigma-B regulator
19	7	1.4	139	2	hypothetical prote
20	7	1.4	141	2	pathogenesis-relat
21	7	1.4	143	2	truncated pili NM
22	7	1.4	147	1	hypothetical prote
23	7	1.4	149	2	pathogenesis-relat
24	7	1.4	161	2	thymocyte protein
25	7	1.4	164	2	hypothetical prote
26	7	1.4	164	2	ethanolamine utili
27	7	1.4	165	2	OV25-2 protein - n
28	7	1.4	173	2	OV25-1 protein - n
29	7	1.4	177	2	small heat shock p

30	7	1.4	178	2	D71451	MG0653 homolog PHO
31	7	1.4	179	2	S29691	AV25 protein - nem
32	7	1.4	179	2	A75215	MG0653 homolog PAB
33	7	1.4	196	2	I39698	blue copper-bindin
34	7	1.4	196	2	TS1838	blue copper bindin
35	7	1.4	196	2	S41270	ribosomal protein
36	7	1.4	201	2	G71428	hypothetical prote
37	7	1.4	206	2	AC2301	hypothetical prote
38	7	1.4	217	2	T40730	probable rna-bindi
39	7	1.4	221	2	E86425	hypothetical prote
40	7	1.4	233	1	C39142	mrxB protein precu
41	7	1.4	237	1	E70982	probable magnesium
42	7	1.4	237	1	S73990	ribosomal protein
43	7	1.4	249	2	F75272	N-acetylglutamate
44	7	1.4	253	2	F81440	probable short-cha
45	7	1.4	257	2	T48058	RING-H2 zinc flinge

ALIGNMENTS

RESULT 1

S40330
Ig kappa chain V-J region - human
C/Species: Homo sapiens (man)
C/Date: 19-May-1994 #sequence_revision 26-May-1995 #text_change 23-Jul-1999
C/Accession: S40330
R/Klein, R.; Jaenichen, R.; Zachau, H.G.
Eur. J. Immunol. 23, 3248-3271, 1993
A/Title: Expressed human immunoglobulin chi genes and their hypermutation.
A/Reference number: S40312; MUID:94080891; PMID:8258341
A/Accession: S40330
A/Status: preliminary; translation not shown
A/Molecule type: mRNA
A/Residues: 1-108 <KLE>
A/Cross-references: EMBL:X72440; NID:9441348; PIDN:CAA51108.1; PID:9441349
A/Superfamily: immunoglobulin V region; immunoglobulin homology
C/Keywords: heterotetramer; immunoglobulin

Query Match 1.7%; Score 9; DB 2; Length 108;

Best Local Similarity 100.0%; Pred. No. 0.44;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 441 SLSVSVD 449
Db 12 SLSVSVD 20

RESULT 2

B95110
competence protein Cella [imported] - Streptococcus pneumoniae (strain TIGR4)
C/Species: Streptococcus pneumoniae
C/Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
C/Accession: B95110
R/Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzaple, I
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A/Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A/Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A/Reference number: A95000; MUID:21357209; PMID:11463916
A/Accession: B95110
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-216 <KUR>
A/Cross-references: GB:AE005672; PIDN:AAK75075.1; PID:G14972428; GSPDB:GN00164; TIGR:SP48
A/Experimental source: strain TIGR4
C/Genetics:
A/Gene: SP0954

Query Match 1.7%; Score 9; DB 2; Length 216;
Best Local Similarity 100.0%; Pred. No. 0.82;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 135 LKXSGIGG 143
 DB 195 LKXSGIGG 203

RESULT 3

H97978
 Competence protein (imported) - Streptococcus pneumoniae (strain R6)
 C:Species: Streptococcus pneumoniae
 C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 22-Oct-2001
 C:Accession: J97978
 R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; Dehoff, B.S.; Ee, R.; LeBlanc, D.J.; Lee, L.N.; Letkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; Meyer, P.; Sun, P.M.; Winkler, M.E.
 J. Bacteriol. 183, 5709-5717, 2001
 A:Author: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaekunas, S.R.; A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
 A:Reference number: A97872; MUID:21429245; PMID:11544234
 A:Accession: H97978
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-216 <KUR>
 A:Cross-references: GB:A807317; PIDN:AAK39660.1; PID:g15458459; GSPDB:GN00174
 C:Genetics:
 A:Gene: cea

Query Match 1.7%; Score 9; DB 2; Length 216;
 Best Local Similarity 100.0%; Pred. No. 0.82;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 135 LKXSGIGG 143
 DB 195 LKXSGIGG 203

RESULT 4

T48159
 Hypothetical protein T1008.80 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
 C:Accession: T48159
 R:Byan, M.; Pohl, T.; Weizsaecker, T.; Bancroft, I.; Mewes, H.W.; Lemcke, K.; Mayer, K. submitted to the Protein Sequence Database, March 2000
 A:Reference number: Z24486
 A:Accession: T48159
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-427 <BEV>
 A:Cross-references: EMBL:AL161746
 A:Experimental source: cultivar Columbia; BAC clone T1008
 C:Genetics:
 A:Map position: 5
 A:Introns: 317/1
 A>Note: T1008.80

Query Match 1.5%; Score 8; DB 2; Length 427;
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 49 SVSFREIV 56
 DB 145 SVSFREIV 152

RESULT 5

A32422
 dihydroliposamide S-(2-methylpropenyl)-transferase (EC 2.3.1.-) precursor - human
 N:Alternate names: 3-methyl-2-oxobutanoate dehydrogenase (liposamide) (EC 1.2.4.4) complete
 C:Species: Homo sapiens (man)
 C:Date: 25-Sep-1989 #sequence_revision 19-Oct-1995 #text_change 18-Jun-1999
 C:Accession: A32422; A28655; A28707; A45126; S27194; S50200; S22865
 R:Nobukuni, Y.; Mitsubuchi, H.; Endo, F.; Matsuda, I.

Biochem. Biophys. Res. Commun. 161, 1035-1041, 1989
 A:Title: Complete primary structure of the transacylase (E2b) subunit of the human branched-chain acyltransferase
 A:Reference number: A32422; MUID:89302075; PMID:2742576
 A:Accession: A32422
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 6-482 <NOB>
 A:Cross-references: GB:M27093; NID:g736674; PIDN:AAA64512.1; PID:g736675
 R:Danner, D.J.; Litwer, S.; Herring, W.J.; Pruckler, J.
 J. Biol. Chem. 264, 7742-7746, 1989
 A:Title: Construction and nucleotide sequence of a cDNA encoding the full-length prepro-
 A:Reference number: A33362; MUID:89214230; PMID:2708389
 A:Accession: A33362
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 6-320, 'O', 322-353, 'L', 355-482 <DAN>
 A:Cross-references: GB:J03208; NID:g19355; PIDN:AAA5589.1; PID:g179354; GB:J04723
 R:Hummel, K.B.; Litwer, S.; Bradford, A.P.; Aitken, A.; Danner, D.J.; Yeaman, S.J.
 J. Biol. Chem. 263, 6165-6168, 1988
 A:Title: Nucleotide sequence of a cDNA for branched chain acyltransferase with analysis
 A:Reference number: A28655; MUID:88198156; PMID:3245861
 A:Accession: A28655
 A:Molecule type: mRNA
 A:Residues: 6-313, 'YFSPWKS', <HUM>
 A:Cross-references: GB:J03208; GB:J04723
 R:Lau, K.S.; Griffin, T.A.; Hu, C.W.C.; Chuang, D.T.
 Biochemistry 27, 1972-1981, 1988
 A:Title: Conservation of primary structure in the lipoyl-bearing and dihydrolipoyl dehydro-
 A:Reference number: A90529; MUID:88241022; PMID:2832277
 A:Accession: A28707
 A:Molecule type: mRNA
 A:Residues: 6-313, 'YFSPWKS', <LAN>
 A:Cross-references: GB:M19301; NID:g16635; PIDN:AAA59200.1; PID:g307078
 R:Lau, K.S.; Herring, W.J.; Chuang, J.L.; McKean, M.; Danner, D.J.; Cox, R.P.; Chuang, D
 J. Biol. Chem. 267, 24090-24096, 1992
 A:Title: Structure of the gene encoding dihydrolipoyl transacylase (E2) component of human
 A:Reference number: A45126; MUID:93054781; PMID:11429740
 A:Accession: A45126
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-315-20; 56-61; 82-87; 142-148; 184-188; 255-260; 331-336; 337-342; 401-406; 425-430
 A:Cross-references: GB:X68104
 R:Lau, K.S.; Chuang, J.L.; Herring, W.J.; Danner, D.J.; Cox, R.P.; Chuang, D.T.
 Biochim. Biophys. Acta 1132, 319-321, 1992
 A:Title: The complete cDNA sequence for dihydrolipoyl transacylase (E2) of human branched
 A:Reference number: S27194; MUID:93041936; PMID:1420314
 A:Accession: S27194
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-320, 'O', 322-353, 'L', 355-482 <LAN>
 A:Cross-references: EMBL:X66785; NID:930489; PIDN:CMA47285.1; PID:g30490
 R:Wynn, R.M.; Kochi, H.; Cox, R.P.; Chuang, D.T.
 Biochim. Biophys. Acta 1201, 125-128, 1994
 A:Title: Differential processing of human and rat E1-alpha precursors of the branched-chain
 A:Reference number: S50200; MUID:95002090; PMID:7918575
 A:Accession: S50200
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 47-81 <WYN>
 C:Comment: Deficiency of various subunits of the branched-chain alpha-keto acid dehydrog
 C:Genetics:
 A:Gene: GDB:DBT
 A:Cross-references: GDB:118784; OMIM:248610
 A:Map position: 1p31-1p31
 C:Superfamily: dihydroliposamide acetyltransferase; lipoyl/biotin-binding homology
 C:Keywords: acyltransferase; blocked amino end; mitochondrion; oxidoreductase
 F:1-61/Domain: transit peptide (mitochondrion) #status predicted <SIG>
 F:62-482/Product: liposamide acetyltransferase #status predicted <MAT>
 F:66-139/Domain: lipoyl/biotin-binding homology <LDB>
 F:105/Binding site: liposamide (lys) (covalent) #status predicted
 F:452,456/Active site: His, Asp #status predicted

Query Match 1.5%; Score 8; DB 2; Length 482;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 LPSPKVE 336
 |||||
 Db 217 LPSPKVE 224

RESULT 6

AD1223
 [hypotheoretical protein lmo1188 [imported] - Listeria monocytogenes (strain BGD-e)]

C/Species: Listeria monocytogenes
 C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
 C/Accession: AD1223

R/Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Jones, L.M.; Karet, U.;
 Science 294, 849-852, 2001
 A/Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapkac, G.; Madueno, E.; Maitournam, A.; Meok, C.; Schluecker, T.; Simoes, N.; Tietzer, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A.; Title: Comparative genomes of Listeria species.
 A/Reference number: AB1077; MUID:21537279; PMID:11679669
 A/Accession: AD1223
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-483 <GUA>
 A/Cross-references: GB:NC_003210; PIDN:CMC9266.1; PID:G16410604; GSPDB:GN00177
 A/Experimental source: strain BGD-e
 C/Genetics:
 A/Gene: lmo1188

Query Match 1.5%; Score 8; DB 2; Length 483;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 ITAEQVK 180
 |||||
 Db 390 ITAEQVK 397

RESULT 7

A34283
 H+-transporting two-sector ATPase (EC 3.6.3.14) alpha chain - Methanosarcina barkeri

C/Species: Methanosarcina barkeri
 C/Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 03-Jun-2002
 C/Accession: A34283

R/Inatomi, K.I.; Eya, S.; Maeda, M.; Futai, M.
 J. Biol. Chem. 264, 10954-10959, 1989
 A/Title: Amino acid sequence of the alpha and beta subunits of Methanosarcina barkeri AT
 A/Reference number: A34283; MUID:89291829; PMID:2544575
 A/Accession: A34283

A/Status: not compared with conceptual translation
 A/Molecule type: DNA
 A/Residues: 1-578 <GNA>
 A/Cross-references: GB:J04836; NID:G149819; PIDN:AAA72215.1; PID:G149820

C/Superfamily: vacuolar H+-transporting ATPase 69k chain; H+-transporting ATP synthase
 C/Keywords: ATP biosynthesis; hydrolase; nucleotide binding; P-loop
 F/228-235/Region: nucleotide-binding motif A (P-loop)
 F/251-431/Domain: H+-transporting ATP synthase alpha chain homology <ATP>

Query Match 1.5%; Score 8; DB 2; Length 578;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 438 STGSSLSV 445
 |||||
 Db 66 STGSSLSV 73

RESULT 8

122340
 [hypotheoretical protein F53b6.6 - Caenorhabditis elegans]

C/Species: Caenorhabditis elegans
 C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C/Accession: T22540
 R/White, S.

submitted to the EMBL Data Library, October 1996
 A/Reference number: Z19578
 A/Accession: T22540

A/Status: preliminary; translated from GB/EMBL/DDBJ
 A/Molecule type: DNA
 A/Residues: 1-610 <WTL>

A/Cross-references: EMBL:Z61086; PIDN:CAM03116.1; GSPDB:GN00019; CESP:F53B6.6
 A/Experimental source: clone F53B6
 C/Genetics:
 A/Gene: CESP:F53B6.6
 A/Map position: 1

A/Insertions: 38/1; 103/1; 192/2; 239/1; 268/2; 300/3; 532/2; 546/1

Query Match 1.5%; Score 8; DB 2; Length 610;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 286 SESSTPST 293
 |||||
 Db 481 SESSTPST 488

RESULT 9

E82768
 conserved hypothetical protein XF0752 [imported] - Xylella fastidiosa (strain 945C)

C/Species: Xylella fastidiosa
 C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
 C/Accession: E82768

R/anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing
 Nature 406, 151-157, 2000
 A/Title: The genome sequence of the plant pathogen Xylella fastidiosa.
 A/Reference number: A82515; MUID:20365717; PMID:10910347
 A/Note: for a complete list of authors see reference number A59328 below

A/Accession: E82768
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-621 <SIM>

A/Cross-references: GB:AE003916; GB:AB003849; NID:99105626; PIDN:AAF83562.1; GSPDB:GN001;
 A/Experimental source: strain 945C
 R/Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A) Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H

as-Heto, E.; Docena, C.; El-Dorty, H.; Facinanti, A.P.; Ferreira, A.J.S.
 submitted to GenBank, June 2000
 A/Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franco, S.C.; Franco, M.C.; Frohne

J.D.; Junqueira, M.L.; Kemper, E.L.; Kitejima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigret
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
 A/Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;

F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmitieri, D.A
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
 A/Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
 M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z

A/Reference number: A59328
 A/Contents: annotation
 C/Genetics:
 A/Gene: XF0752

Query Match 1.5%; Score 8; DB 2; Length 621;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 445 VSVYDAEA 452
 |||||
 Db 452 VSVYDAEA 459

RESULT 10

D75598
 photoreceptor - Deinococcus radiodurans (strain R1)

C/Species: Deinococcus radiodurans

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Feb-2003
C/Accession: D75598
R/White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; Ma
Science 286, 1571-1577, 1999
A>Title: Genome sequence of the radioreistant bacterium Deinococcus radiodurans R1.
A/Reference number: A75250; MUID:20036896; PMID:10567266
A/Accession: D75598
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-755 <WHI-
A/Cross-references: GB:AE001862; GB:AE001825; NID:g6460468; PIDN:AAF12261.1; PID:g646055
A/Experimental source: strain R1
C/Genetics:
A/Gene: DRA0050
A/Map position: 2
C/Superfamily: bacteriophytochrome; phytochrome homology
F:504-744/Domain: sensor histidine kinase homology <SHK>

Query Match 1.5%; Score 8; DB 2; Length 755;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 440 GSSLSVSV 447
DB 270 GSSLSVSV 277

RESULT 11
H96795
hypothetical protein F28016.11 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C/Accession: H96795
R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luross, J.S.; Malt, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yi, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A66141; MUID:21016719; PMID:11130712
A/Accession: H96795
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1532 <SNO>
A/Cross-references: GB:AE005173; NID:g6143891; PIDN:AAF04437.1; GSPDB:GN00141
C/Genetics:
A/Gene: F28016.11
A/Map position: 1

Query Match 1.5%; Score 8; DB 2; Length 1532;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 435 RSGSTGSS 442
DB 146 RSGSTGSS 153

RESULT 12
H91246
Probable DNA-binding protein [imported] - Escherichia coli (strain O157:H7, substrain R1
C/Species: Escherichia coli
C/Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 17-Mar-2003
C/Accession: H91246
R/Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shibata, T.; Hattori, M.; Shingawa, H.
DNA Res. 8, 11-22, 2001

A>Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A/Reference number: A99629; MUID:21156231; PMID:11258796
A/Accession: H91246
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-82 <HAY-
A/Cross-references: GB:BA000007; PIDN:BA38367.1; PID:g13364420; GSPDB:GN00154
A/Experimental source: strain O157:H7, substrain R1MD 050952
C/Genetics:
A/Gene: ECs4944
C/Superfamily: phage D108 DNA-binding protein

Query Match 1.4%; Score 7; DB 2; Length 82;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 317 SLSLKAG 323
DB 29 SLSLKAG 35

RESULT 13
C95304
hypothetical transposase, partial match [imported] - Sinorhizobium meliloti (strain 1021)
C/Species: Sinorhizobium meliloti
C/Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C/Accession: C95304
R/Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bowe
proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A>Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti
A/Reference number: A95262; MUID:21396509; PMID:11481432
A/Accession: C95304
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-86 <KUR-
A/Cross-references: GB:AE006469; PIDN:AAK64997.1; PID:g14523425; GSPDB:GN00165
A/Experimental source: strain 1021, megaplasmid pSymA
R/Galibert, F.; Fhan, T.W.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A/Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yen, K.C.
A>Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A/Reference number: A96039; MUID:21368234; PMID:11474104
A/Contents: annotation
C/Genetics:
A/Gene: Sma0643
A/Genome: plasmid

Query Match 1.4%; Score 7; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 423 AGPELLR 429
DB 66 AGPELLR 72

RESULT 14
KH10UX
Ig kappa chain V-1 region (Iray) - human
C/Species: Homo sapiens (man)
C/Date: 24-Apr-1984 #sequence_revision 24-Apr-1984 #text_change 21-Jan-2000
C/Accession: A01871; F30609
R/Capra, J.D.; Klapper, D.G.
Scand. J. Immunol. 5, 677-684, 1976
A>Title: Complete amino acid sequence of the variable domains of two human IgM anti-gamma
A/Reference number: A01871; MUID:77038198; PMID:824717
A/Accession: A01871
A/Molecule type: protein
A/Residues: 1-108 <CAP>

A;Note: the second and third hypervariable regions of this chain are identical with those
 R;Goni, F.R.; Chen, P.P.; McGinnis, D.; Arjonilla, M.L.; Fernandez, J.; Carson, D.; Solid
 J. Immunol. 142, 3158-3163, 1989
 A;Title: Structural and idiotypic characterization of the L chains of human IGM autoantibody
 A;Reference number: A30601; MUID:89215279; PMID:2496160
 A;Accession: F30609
 A;Molecule type: protein
 A;Residues: 1-104 <GON>
 C;Comment: This chain was isolated from an IGM with anti-gamma globulin activity.
 C;Genetics:
 A;Gene: GDB:IGKV1
 A;Cross-references: GDB:136264
 A;Map position: 2p12-2p12
 C;Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa)
 chain disulfide bonds. In some cases, such as IgA and IGM, the subunits associate into 1a
 C;Superfamily: immunoglobulin V region; immunoglobulin homology
 C;Keywords: heterotetramer; immunoglobulin
 F;16-90/Domain: immunoglobulin homology <IMM>
 F;23-88/Disulfide bonds: #status predicted

Query Match

1.4%; Score 7; DB 1; Length 108;

Best Local Similarity 100.0%; Pred. No. 46;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

441 SLSVSV 447
 |||||

Db 9 SLSVSV 15

RESULT 15

H71014
 hypothetical protein PH1414 - Pyrococcus horikoshii

C;Species: Pyrococcus horikoshii

C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 05-Nov-1999

C;Accession: H71014

R;Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin
 M.; Ohnuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi

DNA Res. 5, 55-76, 1998

A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
 A;Reference number: A71000; MUID:98344137; PMID:9679194

A;Accession: H71014

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-129 <KAW>

A;Cross-references: GB:AP00006; NID:G3236133; PIDN:BAA30520.1; PID:d1031463; PID:g32578

A;Experimental source: strain OT3

A;Note: this accession replaces an interim accession for a sequence replaced by GenBank

C;Genetics:
 A;Gene: PH1414

Query Match

1.4%; Score 7; DB 2; Length 129;

Best Local Similarity 100.0%; Pred. No. 54;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

259 MHLPLEN 265
 |||||

Db 41 MHLPLEN 47

Search completed: July 24, 2003, 12:27:11
 Job time : 64.1903 secs

THIS PAGE BLANK (000710)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 11:44:42 ; Search time 168.369 Seconds
(without alignments)
488.334 Million cell updates/sec

Title: US-09-991-681-27
Perfect score: 2698
Sequence: 1 RIRAMAQQVMDLTQCSPKT.....VROAVREWLGRVGVYDIIV 518

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2698	100.0	518	19	AAW85472
2	2698	100.0	518	23	AAO19165
3	2698	100.0	518	23	AAW50809
4	2698	100.0	1807	22	ABG09728
5	2698	100.0	1839	23	ABG64835
6	2588	95.9	1982	22	ABG09731
7	982	36.4	192	20	AA18248
8	920	34.1	180	20	AA18216
9	260	9.6	49	23	AAO19169

10	260	9.6	49	23	AAW50813	PS118 prostate mar
11	255	9.5	50	22	ABG09729	Novel human diagno
12	216	8.0	41	23	AAO19166	Human prostate-spe
13	216	8.0	41	23	AAW50810	PS118 prostate mar
14	215	8.0	40	23	AAO19168	Human prostate-spe
15	215	8.0	40	23	AAW50812	PS118 prostate mar
16	180	6.7	35	23	AAO19167	Human prostate-spe
17	180	6.7	35	23	AAW50811	PS118 prostate mar
18	177.5	6.6	2045	22	ABG61941	Drosophila melanog
19	133	4.9	665	22	ABG64312	Drosophila melanog
20	122	4.5	1657	22	ABG72031	Drosophila melanog
21	115.5	4.3	342	24	ABR01800	Human cancer-relat
22	113.5	4.2	743	22	AAE09851	Novel human protei
23	113.5	4.2	877	22	AAE09850	Novel human protei
24	113.5	4.2	908	22	AAE09849	Novel human protei
25	113.5	4.2	909	22	AAE09845	Novel human protei
26	113.5	4.2	1042	22	AAE09838	Novel human protei
27	113.5	4.2	1043	22	AAE09844	Novel human protei
28	111	4.1	474	18	AAW15253	Human brain-specif
29	111	4.1	594	18	AAW15256	Human brain-specif
30	111	4.1	1148	20	AAV07087	Renal cancer assoc
31	111	4.1	1189	22	AAU32012	Novel human secret
32	110	4.1	1295	24	ABR01846	Human cancer-relat
33	110	4.1	792	22	AAE09853	Novel human protei
34	110	4.1	827	22	AAE09855	Novel human protei
35	110	4.1	829	13	ABG27549	MCC gene product.
36	110	4.1	829	23	ABG95120	Human translocatio
37	110	4.1	926	22	AAE09852	Novel human protei
38	110	4.1	957	22	AAE09841	Novel human protei
39	110	4.1	958	22	AAE09847	Novel human protei
40	110	4.1	961	22	AAE09854	Novel human protei
41	110	4.1	992	22	AAE09843	Novel human protei
42	110	4.1	993	22	AAE09849	Novel human protei
43	110	4.1	1091	22	AAE09840	Novel human protei
44	110	4.1	1092	22	AAE09846	Novel human protei
45	110	4.1	1126	22	AAE09842	Novel human protei

ALIGNMENTS

RESULT 1	AAW85472	standard; Protein; 518 AA.
ID	AAW85472	
AC	AAW85472;	
XX		
DT	25-MAR-2003	(updated)
DT	25-FEB-1999	(first entry)
XX		
DE	PS118 protein encoded by consensus sequence.	
XX		
KW	EST clone; PS118; prostate tumour tissue; prostatic disease; cancer.	
XX		
OS	Homo sapiens.	
XX		
PN	WO9848054-A1.	
XX		
PD	29-OCT-1998.	
XX		
PF	23-APR-1998;	98MO-US08239.
XX		
PR	23-APR-1997;	97US-0842385.
XX		
PA	(ABBO) ABBOTT LAB.	
XX		
PI	Billing-medcl PA, Cohen M, Colpites TL, Friedman PN, Gordon J,	
PI	Granados EN, Hodges SC, Klass MR, Kratochvil JD, Robertstrapp L,	
PI	Russell JC, Stroupe SD;	
XX		
DR	WPI; 1998-610000/51.	
DR	N-PSDB; AAW82812.	
XX		

PT New P118 nucleic acid and proteins - used for diagnosis and
PT treatment of prostatic disease, especially cancer, and also for drug
screening

Claim 17; Page 93-94; 117pp; English.

CC The present sequence is encoded by consensus P5118 sequence derived from
CC overlapping EST (expressed sequence tag) clones AAV82803-10. The clones
CC were identified from a cDNA library made from prostate tumour tissue.
CC Recombinant P5118 protein is used to detect P5118-specific antibodies,
CC to raise antibodies for detection of P5118 antigens, to screen for
CC specific binding agents (potential therapeutics, and to isolate specific
CC antibodies from serum. Detection of P5118 protein or nucleic acid, which
CC are prostate related, and altered or elevated in prostatic disease, is
CC used for detection, diagnosis, staging, monitoring and prognosis of
CC prostatic disease, particularly cancer, and to identify subjects at
CC risk.

(Updated on 25-MAR-2003 to correct PI field.)

CC Sequence 518 AA;

Query Match 100.0%; Score 2698; DB 19; Length 518;
Best Local Similarity 100.0%; Pred. No. 7.3e-254;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 RIRAAQOVFMDLTQCSPTKPNNDHAQSCQIILPDEKNGHTKXSVFRIIVSL 60
DB 1 RIRAAQOVFMDLTQCSPTKPNNDHAQSCQIILPDEKNGHTKXSVFRIIVSL 60
QY 61 SHQVLTQNLVDLLEEFVKGPSPEEKTIOVPEAKLAGFRIRYSQMOLAVIFDILLDSYR 120
DB 61 SHQVLTQNLVDLLEEFVKGPSPEEKTIOVPEAKLAGFRIRYSQMOLAVIFDILLDSYR 120
QY 121 TARFDTSPGKLCILKVKVSGIGGANLYROSAMSFNIYFHALVCAVLTNOETITAEQVKK 180
DB 121 TARFDTSPGKLCILKVKVSGIGGANLYROSAMSFNIYFHALVCAVLTNOETITAEQVKK 180
QY 121 TARFDTSPGKLCILKVKVSGIGGANLYROSAMSFNIYFHALVCAVLTNOETITAEQVKK 180
DB 121 TARFDTSPGKLCILKVKVSGIGGANLYROSAMSFNIYFHALVCAVLTNOETITAEQVKK 180
QY 181 VLFEDDERSTDSQCCSEDEDIFEETAQVSPPRKEKORAMRPLISVQPVSNADWV 240
DB 181 VLFEDDERSTDSQCCSEDEDIFEETAQVSPPRKEKORAMRPLISVQPVSNADWV 240
QY 181 VLFEDDERSTDSQCCSEDEDIFEETAQVSPPRKEKORAMRPLISVQPVSNADWV 240
DB 181 VLFEDDERSTDSQCCSEDEDIFEETAQVSPPRKEKORAMRPLISVQPVSNADWV 240
QY 241 LVKRLHKLCELCNNYIQMHLDLENCEBPPIFKGDPFFILPSFQSSSTPSTGFSGKE 300
DB 241 LVKRLHKLCELCNNYIQMHLDLENCEBPPIFKGDPFFILPSFQSSSTPSTGFSGKE 300
QY 241 LVKRLHKLCELCNNYIQMHLDLENCEBPPIFKGDPFFILPSFQSSSTPSTGFSGKE 300
DB 241 LVKRLHKLCELCNNYIQMHLDLENCEBPPIFKGDPFFILPSFQSSSTPSTGFSGKE 300
QY 301 TPSEDDRSQSRHEHGESLSLAKAGGDLILPPSPVKEKDPKRKEMWENAGNKIYTMAAD 360
DB 301 TPSEDDRSQSRHEHGESLSLAKAGGDLILPPSPVKEKDPKRKEMWENAGNKIYTMAAD 360
QY 361 KTISLMTVEYKRRKQNHLSAPFKEVKYKGEKGPSPGQSPILQROHLMDOQMHS 420
DB 361 KTISLMTVEYKRRKQNHLSAPFKEVKYKGEKGPSPGQSPILQROHLMDOQMHS 420
QY 421 FSAGBELLRQDKRPPSSGTSGLSVSRDAEAQIQAMTNMVLTVANOQIILPDQFTALQ 480
DB 421 FSAGBELLRQDKRPPSSGTSGLSVSRDAEAQIQAMTNMVLTVANOQIILPDQFTALQ 480
QY 481 PAVFPICISQLTCHVTDIRVQAVREWLGRVGRVYDIIV 518
DB 481 PAVFPICISQLTCHVTDIRVQAVREWLGRVGRVYDIIV 518

```

RESULT 2

AA019165 ID AA019165 standard; Protein; 518 AA.

AA019165;

27-NOV-2002 (first entry)

Human prostate-specific P5118 protein fragment #1.

Human; prostate; prostate-specific sequence; prostate cancer; P5118;

KM EST; expressed sequence tag; cytostatic; gene therapy.

OS Homo sapiens.

PN US2002086316-A1.

04-JUL-2002.

26-NOV-2001; 2001US-0991681.

23-APR-1998; 98US-0065383.

23-APR-1997; 97US-0842385.

(BTL)/ BILLINGEL P. A.

(COHE)/ COHEN M.

(COLP)/ COLPITTS T. L.

(FRIE)/ FRIEDMAN P. N.

(GORD)/ GORDAN J.

(GRAN)/ GRANADOS E. N.

(HODG)/ HODGES S. C.

(KLAS)/ KLAS M. R.

(KRAT)/ KRATOCHVIL J. D.

(ROBE)/ ROBERTS-RAPP L.

(RUS)/ RUSSELL J. C.

(STRO)/ STROUPE S. D.

Billingsel P. A., Cohen M., Colpitts T. L., Friedman P. N., Gordan J., Granados E. N., Hodges S. C., KLAS M. R., Kratochvil J. D., Roberts-Rapp L., Russell J. C., Stroupe S. D.

WPI; 2002-665429/71.

Novel P5118 polypeptide for detecting, diagnosing, staging, monitoring,

prognosticating, preventing, treating, or determining predisposition of

individual to diseases and conditions of prostate, e.g. prostate

cancer -

Claim 17; Page 42-43; 58pp; English.

The present invention relates to a number of prostate-specific sequences

derived from the human P5118 gene. These can be used in the detection,

monitoring and treatment of prostate diseases, particularly prostate

cancer. The present sequence is a P5118 protein fragment of the

invention. The coding sequences of the invention were isolated from a

prostate tissue expressed sequence tag (EST) library.

Sequence 518 AA;

```

QY 1 RIRAAQOVFMDLTQCSPTKPNNDHAQSCQIILPDEKNGHTKXSVFRIIVSL 60
DB 1 RIRAAQOVFMDLTQCSPTKPNNDHAQSCQIILPDEKNGHTKXSVFRIIVSL 60
QY 61 SHQVLTQNLVDLLEEFVKGPSPEEKTIOVPEAKLAGFRIRYSQMOLAVIFDILLDSYR 120
DB 61 SHQVLTQNLVDLLEEFVKGPSPEEKTIOVPEAKLAGFRIRYSQMOLAVIFDILLDSYR 120
QY 121 TARFDTSPGKLCILKVKVSGIGGANLYROSAMSFNIYFHALVCAVLTNOETITAEQVKK 180
DB 121 TARFDTSPGKLCILKVKVSGIGGANLYROSAMSFNIYFHALVCAVLTNOETITAEQVKK 180
QY 181 VLFEDDERSTDSQCCSEDEDIFEETAQVSPPRKEKORAMRPLISVQPVSNADWV 240
DB 181 VLFEDDERSTDSQCCSEDEDIFEETAQVSPPRKEKORAMRPLISVQPVSNADWV 240
QY 241 LVKRLHKLCELCNNYIQMHLDLENCEBPPIFKGDPFFILPSFQSSSTPSTGFSGKE 300
DB 241 LVKRLHKLCELCNNYIQMHLDLENCEBPPIFKGDPFFILPSFQSSSTPSTGFSGKE 300
QY 301 TPSEDDRSQSRHEHGESLSLAKAGGDLILPPSPVKEKDPKRKEMWENAGNKIYTMAAD 360

```


PE 30-MAR-2001; 2001WO-US08631.
 XX 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX (HYSE-) HYSEQ INC.
 XX Drmanac RT, Liu C, Tang YT;
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS73915.
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 PS Claim 20; SEQ ID No 40087; 103pp; English.
 XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC XX
 SQ Sequence 1807 AA;
 Query Match 100.0%; Score 2698; DB 22; Length 1807;
 Best Local Similarity 100.0%; Pred. No. 5e-253;
 Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 RIRAMAQQVFMMDTCCSPKTPNNFPHASCOLIELPDEKNGHTKKSVPREIVSL 60
 DB 1290 RIRAMAQQVFMMDTCCSPKTPNNFPHASCOLIELPDEKNGHTKKSVPREIVSL 1349
 Oy 61 SHOVLLOMLYDILLEEYKSPGSEKTIQVEEAKIAGFLRYISQNLAVIFDILLDSYR 120
 DB 1350 SHOVLLOMLYDILLEEYKSPGSEKTIQVEEAKIAGFLRYISQNLAVIFDILLDSYR 1409
 Oy 121 TARBDTSPGKCLIKXVSGIGGANLYRQSMSTNIFHALVCVILNCFITAEQYK 180
 DB 1410 TARBDTSPGKCLIKXVSGIGGANLYRQSMSTNIFHALVCVILNCFITAEQYK 1469
 Oy 181 VLFEDDSTDSQOCCSEDEDIFETAVQSPPRKEKROWRARPLLSVOPVSADAVW 240
 DB 1470 VLFEDDSTDSQOCCSEDEDIFETAVQSPPRKEKROWRARPLLSVOPVSADAVW 1529
 Oy 241 LVKRLHKLCMBLCNNYIOHMLDLNCEMEEPPIFKDPPFILPSFGSSSTPSTGFSKE 300
 DB 1530 LVKRLHKLCMBLCNNYIOHMLDLNCEMEEPPIFKDPPFILPSFGSSSTPSTGFSKE 1589
 Oy 301 TPSEDDRSOSRHHMESLSIKAGGDDLPPSPKYEKDPSPKKEKMMENAGKITYTMAD 360
 DB 1590 TPSEDDRSOSRHHMESLSIKAGGDDLPPSPKYEKDPSPKKEKMMENAGKITYTMAD 1649
 Oy 361 KTISKLTMEYKKRQOHNLSAFPKEVKEKGEPIVGPGRGDSPLLQRFQHLMDGOMRHS 420

DB 1650 KTISKLTMEYKKRQOHNLSAFPKEVKEKGEPIVGPGRGDSPLLQRFQHLMDGOMRHS 1709
 Oy 421 FSAGPELLRODKRPRSGSTSSLSVSRDAEAOIQANTNNVLYNLQIQLPDQFTALQ 480
 DB 1710 FSAGPELLRODKRPRSGSTSSLSVSRDAEAOIQANTNNVLYNLQIQLPDQFTALQ 1769
 Oy 481 PAVFPCISQLTCHTDIRVROAVREMLGRVGRVYDIIV 518
 DB 1770 PAVFPCISQLTCHTDIRVROAVREMLGRVGRVYDIIV 1807
 RESULT 5
 ID ABP64835 standard; Protein; 1839 AA.
 XX ABP64835;
 XX 25-FEB-2003 (first entry)
 DT Human protein SEQ ID 495.
 DE Human, expressed sequence tag; EST;
 XX haematopoietic disorder; central nervous system disease; viral infection;
 KW peripheral nervous system disease; non-healing wound; infectious disease;
 KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;
 KW fungal infection; autoimmune disorder; coagulation disorder; noctropic;
 KW antiallergic; antiinflammatory; immunosuppressive; neuroprotective;
 KW cyostatic; haemostatic; virucide; antibacterial; fungicide;
 KW immunostimulant; cerebroprotective.
 OS Homo sapiens.
 XX WO200259260-A2.
 EN 01-AUG-2002.
 PD 16-NOV-2001; 2001WO-US42950.
 XX 17-NOV-2000; 2000US-0714936.
 PR (HYSE-) HYSEQ INC.
 XX Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;
 PI Ren F, Xue AJ, Yang Y, Weinman T, Drmanac RT;
 XX WPI; 2002-590824/63.
 DR N-PSDB; ABQ99421.
 XX New isolated polynucleotide, useful in research, diagnostic or
 PT therapeutic methods, e.g. preventing or treating disorders involving
 PT aberrant protein expression or biological activity -
 PS Claim 20; SEQ ID 495; 394pp; English.
 XX The present invention relates to novel human coding sequences
 CC (ABQ99268-ABQ99608) and proteins (ABP64682-ABP65022). The sequences are
 CC useful in therapeutic, diagnostic and research methods. The
 CC polynucleotides may be used in the field of molecular biology as
 CC hybridisation probes, primers for PCR, for chromosome and gene mapping,
 CC for the recombinant production of protein, or in generation of anti-gene
 CC DNA or RNA. The polynucleotides are useful in diagnostics as expressed
 CC sequence tags (ESTs) for identifying expressed genes or for physical
 CC mapping of the human genome. The proteins may be used as molecular weight
 CC markers, or as nutritional sources or supplements. The proteins may be
 CC used to maintain and expand cell population in a totipotent or
 CC pluripotent state useful for re-engineering damaged or diseased
 CC tissues, transplantation, manufacture of bio-pharmaceuticals or the
 CC development of bio-sensors. The polynucleotides and proteins are useful
 CC for preventing, treating or ameliorating disorders involving aberrant
 CC protein expression or biological activity, e.g. haematopoietic disorders,
 CC central/peripheral nervous system diseases, mechanical and traumatic
 CC disorders, non-healing wounds, immune deficiencies and disorders,
 CC infectious diseases caused by viral, bacterial or fungal infection.

CC autoimmune disorders, allergic reactions and conditions, coagulation
CC disorders, or cancer. The polynucleotide sequences of the invention were
CC assembled from ESTs isolated mainly by sequencing by hybridisation, and
CC in some cases, sequences obtained from one or more public databases.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1839 AA;

Query Match 100.0%; Score 2698; DB 23; Length 1839;
Best Local Similarity 100.0%; Pred. No. 5.2e-253;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIRMAQOVFMDTQCSPTKPNFPAQSCOLIIELPDEKPNHTKKSVSFEIIVSL 60
DB 1322 RIRMAQOVFMDTQCSPTKPNFPAQSCOLIIELPDEKPNHTKKSVSFEIIVSL 1361
QY 61 SHOVLQNLVDIIEEFVKGPSPEEKTIOVPEAKLAGFLRYISMONLAVIFDILLDSYR 120
DB 1382 SHOVLQNLVDIIEEFVKGPSPEEKTIOVPEAKLAGFLRYISMONLAVIFDILLDSYR 1441
QY 121 TAREFDTSPGLKCLLKVVSGIGGANLYROSAMFNIFYHALVCAVLTNOETTTAEQVKK 180
DB 1442 TAREFDTSPGLKCLLKVVSGIGGANLYROSAMFNIFYHALVCAVLTNOETTTAEQVKK 1501
QY 181 VLFEDDERSTDSQOCSSEDEDI FEETAQVSPRGKEXKQWRAMPPLISVQPSNADVMW 240
DB 1502 VLFEDDERSTDSQOCSSEDEDI FEETAQVSPRGKEXKQWRAMPPLISVQPSNADVMW 1561
QY 241 LVKRLHKLQMLCNVYIOMHLDLNCMEBPPIFKGDPFFILPSFQSSSTPTSGFSGKE 300
DB 1562 LVKRLHKLQMLCNVYIOMHLDLNCMEBPPIFKGDPFFILPSFQSSSTPTSGFSGKE 1621
QY 301 TPSEDDRSQSRHEHGESLSLAKAGGDLILPPSPKVEKDPBRKKEWMENAGNKIYTMAD 360
DB 1622 TPSEDDRSQSRHEHGESLSLAKAGGDLILPPSPKVEKDPBRKKEWMENAGNKIYTMAD 1681
QY 361 KTISKLMTEYKKRQOHNLSAFPKEVKEKGEPLPGPGDPSPLQRPQHLMDQGQWRHS 420
DB 1682 KTISKLMTEYKKRQOHNLSAFPKEVKEKGEPLPGPGDPSPLQRPQHLMDQGQWRHS 1741
QY 421 FSAGPELLRODKRRSGSTGSSLSVSRDAEAQIOAWTNMVLTVLNOIQLPDTFTALQ 480
DB 1742 FSAGPELLRODKRRSGSTGSSLSVSRDAEAQIOAWTNMVLTVLNOIQLPDTFTALQ 1801
QY 481 PAVFPCISQLTCHVTDIRVROAVREBWLGRVGRVYDIIV 518
DB 1802 PAVFPCISQLTCHVTDIRVROAVREBWLGRVGRVYDIIV 1839

RESULT 6

ABG09731 ID ABG09731 standard; Proteiin, 1982 AA.

XX AC ABG09731;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #9722.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

PI Dmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

XX N-PSDB; AAS73918.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

PS Claim 20; SEQ ID No 40090; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1982 AA;

Query Match 95.9%; Score 2588; DB 22; Length 1982;
Best Local Similarity 99.2%; Pred. No. 3.1e-242;
Matches 499; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

QY 1 RIRMAQOVFMDTQCSPTKPNFPAQSCOLIIELPDEKPNHTKKSVSFEIIVSL 60
DB 1408 RIRMAQOVFMDTQCSPTKPNFPAQSCOLIIELPDEKPNHTKKSVSFEIIVSL 1463
QY 61 SHOVLQNLVDIIEEFVKGPSPEEKTIOVPEAKLAGFLRYISMONLAVIFDILLDSYR 120
DB 1464 SHOVLQNLVDIIEEFVKGPSPEEKTIOVPEAKLAGFLRYISMONLAVIFDILLDSYR 1523
QY 121 TAREFDTSPGLKCLLKVVSGIGGANLYROSAMFNIFYHALVCAVLTNOETTTAEQVKK 180
DB 1524 TAREFDTSPGLKCLLKVVSGIGGANLYROSAMFNIFYHALVCAVLTNOETTTAEQVKK 1583
QY 181 VLFEDDERSTDSQOCSSEDEDI FEETAQVSPRGKEXKQWRAMPPLISVQPSNADVMW 240
DB 1584 VLFEDDERSTDSQOCSSEDEDI FEETAQVSPRGKEXKQWRAMPPLISVQPSNADVMW 1643
QY 241 LVKRLHKLQMLCNVYIOMHLDLNCMEBPPIFKGDPFFILPSFQSSSTPTSGFSGKE 300
DB 1644 LVKRLHKLQMLCNVYIOMHLDLNCMEBPPIFKGDPFFILPSFQSSSTPTSGFSGKE 1703
QY 301 TPSEDDRSQSRHEHGESLSLAKAGGDLILPPSPKVEKDPBRKKEWMENAGNKIYTMAD 360
DB 1704 TPSEDDRSQSRHEHGESLSLAKAGGDLILPPSPKVEKDPBRKKEWMENAGNKIYTMAD 1763
QY 361 KTISKLMTEYKKRQOHNLSAFPKEVKEKGEPLPGPGDPSPLQRPQHLMDQGQWRHS 420
DB 1764 KTISKLMTEYKKRQOHNLSAFPKEVKEKGEPLPGPGDPSPLQRPQHLMDQGQWRHS 1823
QY 421 FSAGPELLRODKRRSGSTGSSLSVSRDAEAQIOAWTNMVLTVLNOIQLPDTFTALQ 480

DB 1824 PSAGELRLRQDRKPSGSGSSLSVSDPAEAOIAMIMVLTALNOIILPDQFTALQ 1883

OY 481 PAVFPCISQLTCHVTDIRVQAV 503

DB 1884 PAVFPCISQLTCHVTDIRVQAV 1906

RESULT 7

AAV48248 ID AAV48248 standard; Protein; 192 AA.

AAV48248; AC

DT 08-DEC-1999 (first entry)

DE Human prostate cancer-associated protein 34.

KM Expressed sequence tag; EST; prostate tumor; antitumor; treatment;
XX gene therapy; tissue specificity human.

OS Homo sapiens.

PN DE1981193-A1.

XX 16-SEP-1999.

PF 10-MAR-1998; 98DE-1011193.

PR 10-MAR-1998; 98DE-1011193.

PA (META-) METAGEN GES GENOMFORSCHUNG MBH.

PI Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosenthal A;

DR WPI, 1999-519628/44.

DR N-PSDB; AAZ33451.

PT New nucleic acid expressed at high level in prostatic tumor tissue and
XX encoded polypeptides, useful for treating cancer and screening for
PT therapeutic agents

PS Claim 22; 128; 166p; German.

CC This invention describes novel nucleic acid sequences (A) that are
XX expressed at high level in prostatic tumor tissue and encode gene
CC products or their fragments. The products of the invention have
CC antitumor activity. Polypeptides (I) encoded by (A) are used: (i) for
CC identifying agents for treatment of prostatic cancer and (ii) for
CC therapy of prostate cancer, optionally where expressed by gene therapy
CC method. (A) is also used to isolate full-length genes (for gene therapy)
CC and for recombinant production of (I), which can be used to raise
CC specific antibodies. (A) are identified by assembly of ESTs (expressed
CC sequence tags) before they are analyzed for expression pattern (tissue
CC specificity). This approach eliminates many of the false results, as
CC regards tissue specificity, associated with known methods that use
CC single (usually short) ESTs. AAV48215-Y48303 represent protein fragments
CC encoded by the expressed sequence tags described in the method of the
CC invention.

XX Sequence 192 AA;

Query Match 36.4%; Score 982; DB 20; Length 192;

Best Local Similarity 100.0%; Pred. No. 3.7e-87;

Matches 192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 MAQOVFMDPTQCSPTNNFPAQSCOLIIEPPEKPNKGTGKSVSPREIVSLSHOV 64

DB 1 MAQOVFMDPTQCSPTNNFPAQSCOLIIEPPEKPNKGTGKSVSPREIVSLSHOV 60

OY 65 LLQNLVYDILBEFVKGSPGEKTIQVPEAKLAGFLRYISMQNLAVIFDILLDSYRTARE 124

DB 61 LLQNLVYDILBEFVKGSPGEKTIQVPEAKLAGFLRYISMQNLAVIFDILLDSYRTARE 120

OY 125 FDTSPGLKCLKKXSGIGGAANLYROSAMSPNIFYHALVCAVLTNOETITAEQVKYLF 184

DB 121 FDTSPGLKCLKKXSGIGGAANLYROSAMSPNIFYHALVCAVLTNOETITAEQVKYLF 180

OY 185 DDERSTDSQOC 196

DB 181 DDERSTDSQOC 192

RESULT 8

AAV48216 ID AAV48216 standard; Protein; 180 AA.

AAV48216; AC

DT 08-DEC-1999 (first entry)

DE Human prostate cancer-associated protein 2.

KM Expressed sequence tag; EST; prostate tumor; antitumor; treatment;
XX gene therapy; tissue specificity human.

OS Homo sapiens.

PN DE1981193-A1.

XX 16-SEP-1999.

PF 10-MAR-1998; 98DE-1011193.

PR 10-MAR-1998; 98DE-1011193.

PA (META-) METAGEN GES GENOMFORSCHUNG MBH.

PI Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosenthal A;

DR WPI, 1999-519628/44.

DR N-PSDB; AAZ33424.

PT New nucleic acid expressed at high level in prostatic tumor tissue and
XX encoded polypeptides, useful for treating cancer and screening for
PT therapeutic agents

PS Claim 22; 112-113; 166p; German.

CC This invention describes novel nucleic acid sequences (A) that are
XX expressed at high level in prostatic tumor tissue and encode gene
CC products or their fragments. The products of the invention have
CC antitumor activity. Polypeptides (I) encoded by (A) are used: (i) for
CC identifying agents for treatment of prostatic cancer and (ii) for
CC therapy of prostate cancer, optionally where expressed by gene therapy
CC method. (A) is also used to isolate full-length genes (for gene therapy)
CC and for recombinant production of (I), which can be used to raise
CC specific antibodies. (A) are identified by assembly of ESTs (expressed
CC sequence tags) before they are analyzed for expression pattern (tissue
CC specificity). This approach eliminates many of the false results, as
CC regards tissue specificity, associated with known methods that use
CC single (usually short) ESTs. AAV48215-Y48303 represent protein fragments
CC encoded by the expressed sequence tags described in the method of the
CC invention.

XX Sequence 180 AA;

Query Match 34.1%; Score 920; DB 20; Length 180;

Best Local Similarity 100.0%; Pred. No. 3.7e-81;

Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 MLDTPQCSBKTPNNFPAQSCOLIIEPPEKPNKGTGKSVSPREIVSLSHOVLONLY 70

DB 1 MLDTPQCSBKTPNNFPAQSCOLIIEPPEKPNKGTGKSVSPREIVSLSHOVLONLY 60

OY 71 DILBEFVKGSPGEKTIQVPEAKLAGFLRYISMQNLAVIFDILLDSYRTAREFDTSPG 130

```

Db      61 DILLEEFVKGPSPEBEKTIQVPEAKLAGFLRYISMQNLAVIFDILLDSYRTAREFTSPG 120
Qy      131 LKCLLKRVSGTIGSGAANYROSAMSFNTYFHALVCANVTNQTITAEQVKYLFEEDDERST 190
Db      121 LKCLLKRVSGTIGSGAANYROSAMSFNTYFHALVCANVTNQTITAEQVKYLFEEDDERST 180

RESULT 9
AA019169
ID      AA019169 standard; Protein; 49 AA.
AC      AA019169;
XX
XX      27-NOV-2002 (first entry)
XX
XX      Human prostate-specific PS118 protein fragment #5.
XX
XX      Human; prostate; prostate-specific sequence; prostate cancer; PS118;
XX      EST; expressed sequence tag; cytostatic; gene therapy.
XX
XX      Homo sapiens.
XX
XX      US2002086316-A1.
XX
XX      04-JUL-2002.
XX
XX      26-NOV-2001; 2001US-0991681.
XX
XX      23-APR-1998; 98US-0065383.
XX      23-APR-1997; 97US-0842385.
XX
XX      (BILL/) BILLINGEL P A.
XX      (COHE/) COHEN M.
XX      (COLP/) COLPITTS T L.
XX      (FRIE/) FRIEDMAN P N.
XX      (GORD/) GORDAN J.
XX      (GRAN/) GRANADOS E N.
XX      (HODG/) HODGES S C.
XX      (KLAS/) KLAS M R.
XX      (KRAT/) KRATOCHVIL J D.
XX      (ROBE/) ROBERTS-RAPP L.
XX      (RUSSE/) RUSSELL J C.
XX      (STROU/) STROUPE S D.
XX
XX      Billingsel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J,
XX      Granados EN, Hodges SC, Kلاس MR, Kratochvil JD, Roberts-Rapp L;
XX      Russell JC, Stroupe SD;
XX      WPI; 2002-665429/71.
XX
XX      Novel PS118 polypeptide for detecting, diagnosing, staging, monitoring,
XX      prognosticating, preventing, treating, or determining predisposition of
XX      individual to diseases and conditions of prostate, e.g. prostate
XX      cancer -
XX
XX      Claim 17; Page 44; 58pp; English.
XX
XX      The present invention relates to a number of prostate-specific sequences
XX      derived from the human PS118 gene. These can be used in the detection,
XX      monitoring and treatment of prostate diseases, particularly prostate
XX      cancer. The present sequence is a PS118 protein fragment of the
XX      invention. The coding sequences of the invention were isolated from a
XX      prostate tissue expressed sequence tag (EST) library.
XX
XX      Sequence 49 AA;
XX
XX      Query Match 9.6%; Score 260; DB 23; Length 49;
XX      Best Local Similarity 100.0%; Pred. No. 1,2e-17;
XX      Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX      393 EPICPRGQDSPLLQRPQHLMDQGMRSFSAGPELLRQDKRPRSGSTGS 441
XX

```

```

Db      1 EPICPRGQDSPLLQRPQHLMDQGMRSFSAGPELLRQDKRPRSGSTGS 49
RESULT 10
AAM50813
ID      AAM50813 standard; Protein; 49 AA.
AC      AAM50813;
XX
XX      01-MAY-2002 (first entry)
XX
XX      PS118 prostate marker immunogenic polypeptide.
XX
XX      PS118; prostate; marker; prostate cancer; tumour; metastasis;
XX      benign prostatic hyperplasia; prostatic intraepithelial neoplasia;
XX      prostaticitis; human; diagnosis; therapy; vaccine; immunogen.
XX
XX      Homo sapiens.
XX
XX      US2001055758-A1.
XX
XX      27-DEC-2001.
XX
XX      23-APR-1998; 98US-0065383.
XX      23-APR-1997; 97US-0842385.
XX
XX      (BILL/) BILLING-MEDEL P A.
XX      (COHE/) COHEN M.
XX      (COLP/) COLPITTS T L.
XX      (FRIE/) FRIEDMAN P N.
XX      (GORD/) GORDON J.
XX      (GRAN/) GRANADOS E N.
XX      (HODG/) HODGES S C.
XX      (KLAS/) KLAS M R.
XX      (KRAT/) KRATOCHVIL J D.
XX      (ROBE/) ROBERTS-RAPP L.
XX
XX      Billings-Medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J,
XX      Granados EN, Hodges SC, Kلاس MR, Kratochvil JD, Roberts-Rapp L;
XX      Russell JC, Stroupe SD;
XX      WPI; 2002-187683/24.
XX
XX      Detecting presence of target PS118 polynucleotide in test sample,
XX      useful for detecting, diagnosing, staging, monitoring, prognosticating,
XX      preventing or treating or determining predisposition to prostate
XX      disease -
XX
XX      Claim 17; Page 44; 57pp; English.
XX
XX      The present sequence is that of an immunogenic polypeptide
XX      comprising amino acids 393-441 of human prostate-specific PS118
XX      polypeptide (see AAM50809). A PS118 consensus sequence (see
XX      ABA91651) is found at least 12 times more often in prostate than
XX      in non-prostate tissue. PS118 polypeptides, polynucleotides,
XX      antibodies, agonists and inhibitors are useful for detecting,
XX      diagnosing, staging, monitoring, prognosticating, preventing and
XX      treating (including by genetic immunisation), or determining the
XX      predisposition of an individual to, diseases and conditions of the
XX      prostate, such as benign prostatic hyperplasia, prostaticitis,
XX      prostatic intraepithelial neoplasia, prostate cancer, tumours and
XX      metastases. The PS118 polypeptides can be produced by expression
XX      of PS118 polynucleotides in transfected host cells, and
XX      immunogenic peptides are useful for raising PS118-specific
XX      antibodies of diagnostic use. The methods and reagents of the
XX      invention may provide an early means of detecting diseases of the
XX      prostate and may also provide new markers which can differentiate
XX      between the clinically important and unimportant prostate cancers
XX      without the use of surgery.
XX
XX      Sequence 49 AA;
XX

```

Query Match 9.6%; Score 260; DB 23; Length 49;
 Best Local Similarity 100.0%; Pred. No. 1.2e-17;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

393 EPLGGRGDSPLIQRPOHLMDOGMHRSFASGPELLLRQKRRSGSTGS 441
 1 EPLGGRGDSPLIQRPOHLMDOGMHRSFASGPELLLRQKRRSGSTGS 49

RESULT 11
 ABG09729
 ID ABG09729 standard; Protein; 50 AA.
 XX
 AC ABG09729;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #9720.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI: 2001-639362/73.
 DR N-PSDB; AAS73916.

New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -

Claim 20; SEQ ID No 40088; 103bp; English.

The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostic, forensics, gene mapping, identification of mutations in
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

Sequence 50 AA;
 Query Match 9.5%; Score 255; DB 22; Length 50;
 Best Local Similarity 98.0%; Pred. No. 3.7e-17;

Matches 49; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

460 MVLTVLNOIQILPDQFTALQPAVFPICISQLTCHVTDIRVQAVRWLGR 509
 1 MVLTVLNOIQILPDQFTALQPAVFPICISQLTCHVTDIRVQAVRWLGR 50

RESULT 12
 AA019166
 ID AA019166 standard; Protein; 41 AA.
 XX
 AC AA019166;
 XX
 DT 27-NOV-2002 (first entry)
 XX
 DE Human prostate-specific PS118 protein fragment #2.
 XX
 KW Human; prostate; prostate-specific sequence; prostate cancer; PS118;
 KW EST; expressed sequence tag; cytostatic; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN US2002086316-A1.
 XX
 PD 04-JUL-2002.
 XX
 PF 26-NOV-2001; 2001US-0991681.
 XX
 PR 23-APR-1998; 98US-0065383.
 PR 23-APR-1997; 97US-0842385.
 XX
 PA (BILL/) BILLINGEL P A.
 PA (COHE/) COHEN M.
 PA (COLP/) COLPITTS T L.
 PA (FRIE/) FRIEDMAN P N.
 PA (GORD/) GORDAN J.
 PA (GRAN/) GRANADOS E N.
 PA (HODG/) HODGES S C.
 PA (KLAS/) KLAS M R.
 PA (KRAT/) KRATOCHVIL J D.
 PA (ROBE/) ROBERTS-RAPP L.
 PA (RUS/) RUSSELL J C.
 PA (STRO/) STROUPE S D.

Billingei PA, Cohen M, Colpitts TL, Friedman PN, Gordan J,
 PI Granados EN, Hodges SC, Kلاس MR, Kratochvil JD, Roberts-Rapp L,
 PI Russell JC, Stroupe SD;
 DR WPI: 2002-665429/71.

Novel PS118 polypeptide for detecting, diagnosing, staging, monitoring,
 PT prognosticating, preventing, treating, or determining predisposition of
 PT individual to diseases and conditions of prostate, e.g. prostate
 PT cancer -

Claim 17; Page 43-44; 58bp; English.

The present invention relates to a number of prostate-specific sequences
 CC derived from the human PS118 gene. These can be used in the detection,
 CC monitoring and treatment of prostate diseases, particularly prostate
 CC cancer. The present sequence is a PS118 protein fragment of the
 CC invention. The coding sequences of the invention were isolated from a
 CC prostate tissue expressed sequence tag (EST) library.

Sequence 41 AA;

Query Match 8.0%; Score 216; DB 23; Length 41;
 Best Local Similarity 100.0%; Pred. No. 1.7e-13;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

184 EDDERSTDSQCCSSEDEDFEETAOVSPPGKXKQWRAR 224
 1 EDDERSTDSQCCSSEDEDFEETAOVSPPGKXKQWRAR 41

RESULT 13

AA050810 ID AA050810 standard; Protein; 41 AA.

XX AC AA050810;

XX DT 01-MAY-2002 (first entry)

XX DE PS118 prostate marker immunogenic polypeptide.

XX KM PS118; prostate; marker; prostate cancer; tumour; metastasis;
XX KM benign prostatic hyperplasia; prostatic intraepithelial neoplasia;
XX KM prostatic; human; diagnosis; therapy; vaccine; immunogen.

XX OS Homo sapiens.

XX PN US2001055758-A1.

XX PD 27-DEC-2001.

XX PF 23-APR-1998; 98US-0065383.

XX PR 23-APR-1997; 97US-0842385.

XX PA (BILL/) BILLING-MEDEL P A.

XX PA (COLE/) COHEN M.

XX PA (FRIE/) FRIEDMAN P N.

XX PA (GORD/) GORDAN J.

XX PA (GRAN/) GRANADOS E N.

XX PA (HODG/) HODGES S C.

XX PA (KLAS/) KLAS M R.

XX PA (KRAT/) KRATOCHVIL J D.

XX PA (ROBE/) ROBERTS-RAPP L.

XX PI Billing-Medel PA, Cohen M, Coplitts TL, Friedman PN, Gordon J;
XX PI Granados EN, Hodges SC, KLAS MR, Kratochvil JD, Roberts-Rapp L;
XX PI Russell JC, Stroupe SD;

XX DR WPI; 2002-187683/24.

XX PT Detecting presence of target PS118 polynucleotide in test sample,
XX PT useful for detecting, diagnosing, staging, monitoring, prognosticating,
XX PT preventing or treating or determining predisposition to prostate
XX PT disease -

XX PS Claim 17; Page 43; 57pp; English.

XX CC The present sequence is that of an immunogenic polypeptide
XX CC comprising amino acids 184-224 of human prostate-specific PS118
XX CC polypeptide (see AA050809). A PS118 consensus sequence (see
XX CC AA050810) is found at least 12 times more often in prostate than
XX CC in non-prostate tissue. PS118 polypeptides, polynucleotides,
XX CC antibodies, agonists and inhibitors are useful for detecting,
XX CC diagnosing, staging, monitoring, prognosticating, preventing and
XX CC treating (including by genetic immunisation), or determining the
XX CC predisposition of an individual to, diseases and conditions of the
XX CC prostate, such as benign prostatic hyperplasia, prostatitis,
XX CC prostatic intraepithelial neoplasia, prostate cancer, tumours and
XX CC metastases. The PS118 polypeptides can be produced by expression
XX CC of PS118 polynucleotides in transfected host cells, and
XX CC immunogenic peptides are useful for raising PS118-specific
XX CC antibodies of diagnostic use. The methods and reagents of the
XX CC invention may provide an early means of detecting diseases of the
XX CC prostate and may also provide new markers which can differentiate
XX CC between the clinically important and unimportant prostate cancers
XX CC without the use of surgery.

XX SQ Sequence 41 AA;

Query Match 8.0%; Score 216; DB 23; Length 41;

Best Local Similarity 100.0%; Pred. No. 1.7e-13;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 184 EDDERSTDSQCCSDEDEFEETAOVSPRGKXQWRAR 224

DB 1 EDDERSTDSQCCSDEDEFEETAOVSPRGKXQWRAR 41

RESULT 14

AA019168 ID AA019168 standard; Protein; 40 AA.

XX AC AA019168;

XX DT 27-NOV-2002 (first entry)

XX DE Human prostate-specific PS118 protein fragment #4.

XX KM Human; prostate; prostate-specific sequence; prostate cancer; PS118;
XX KM EST; expressed sequence tag; cytosolic; gene therapy.

XX OS Homo sapiens.

XX PN US2002086316-A1.

XX PD 04-JUL-2002.

XX PF 26-NOV-2001; 2001US-0991681.

XX PR 23-APR-1998; 98US-0065383.

XX PR 23-APR-1997; 97US-0842385.

XX PA (BILL/) BILLINGEL P A.

XX PA (COLE/) COHEN M.

XX PA (FRIE/) FRIEDMAN P N.

XX PA (GORD/) GORDAN J.

XX PA (GRAN/) GRANADOS E N.

XX PA (HODG/) HODGES S C.

XX PA (KLAS/) KLAS M R.

XX PA (KRAT/) KRATOCHVIL J D.

XX PA (ROBE/) ROBERTS-RAPP L.

XX PA (RUS/) RUSSELL J C.

XX PA (STRO/) STROUPE S D.

XX PI Billingel PA, Cohen M, Coplitts TL, Friedman PN, Gordon J;
XX PI Granados EN, Hodges SC, KLAS MR, Kratochvil JD, Roberts-Rapp L;
XX PI Russell JC, Stroupe SD;

XX DR WPI; 2002-665429/71.

XX PT Novel PS118 polypeptide for detecting, diagnosing, staging, monitoring,
XX PT prognosticating, preventing, treating, or determining predisposition of
XX PT individual to diseases and conditions of prostate, e.g. prostate
XX PT cancer -

XX PS Claim 17; Page 44; 58pp; English.

XX CC The present invention relates to a number of prostate-specific sequences
XX CC derived from the human PS118 gene. These can be used in the detection,
XX CC monitoring and treatment of prostate diseases, particularly prostate
XX CC cancer. The present sequence is a PS118 protein fragment of the
XX CC invention. The coding sequences of the invention were isolated from a
XX CC prostate tissue expressed sequence tag (EST) library.

XX SQ Sequence 40 AA;

Query Match 8.0%; Score 215; DB 23; Length 40;
Best Local Similarity 100.0%; Pred. No. 2.1e-13;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 332 SPVVEKKDPSRKKEWENAGNKIYTAADTKISKMTYK 371

|||||

DB 1 SPKVEKDPSPRKKEMWENAGNKIYTMADKTIISKLTMEYK 40

RESULT 15

AAM50812
ID AAM50812 standard; Protein; 40 AA.

XX AAM50812;

DT 01-MAY-2002 (first entry)

DE PS118 prostate marker immunogenic polypeptide.

KM PS118; prostate; marker; prostate cancer; tumour; metastasis;
KM benign prostatic hyperplasia; prostatic intraepithelial neoplasia;
KM prostatitis; human; diagnosis; therapy; vaccine; immunogen.

OS Homo sapiens.

XX US2001055758-A1.

XX 27-DEC-2001.

XX 23-APR-1998; 98US-0065383.

XX 23-APR-1997; 97US-0842385.

PA (BILL/) BILLING-MEDEL P A.

PA (COHE/) COHEN M.

PA (COPL/) COPLITTIS T L.

PA (FRIE/) FRIEDMAN P N.

PA (GORD/) GORDON J.

PA (GRAN/) GRANADOS E N.

PA (HODG/) HODGES S C.

PA (KLAS/) KLAS M R.

PA (KRAT/) KRATOCHVIL J D.

PA (ROBE/) ROBERTS-RAPP L.

PI Billing-Medel PA, Cohen M, Coplittis TU, Friedman PN, Gordon J;

PI Granados EN, Hodges SC, Klass MR, Kratochvil JD, Roberts-Rapp L;

PI Russell JC, Stroupe SD;

DR WPI: 2002-187683/24.

PT Detecting presence of target PS118 polynucleotide in test sample,

PT useful for detecting, diagnosing, staging, monitoring, prognosticating,

PT preventing or treating or determining predisposition to prostate

PT disease -

PS Claim 17; Page 44; 57pp; English.

XX The present sequence is that of an immunogenic polypeptide

CC comprising amino acids 332-371 of human prostate-specific PS118

CC polypeptide (see AAM50809). A PS118 consensus sequence (see

CC ABA91651) is found at least 12 times more often in prostate than

CC in non-prostate tissue. PS118 polypeptides, polynucleotides,

CC antibodies, agonists and inhibitors are useful for detecting,

CC diagnosing, staging, monitoring, prognosticating, preventing and

CC treating (including by genetic immunisation), or determining the

CC predisposition of an individual to, diseases and conditions of the

CC prostate, such as benign prostatic hyperplasia, prostatitis,

CC prostatic intraepithelial neoplasia, prostate cancer, tumours and

CC metastases. The PS118 polypeptides can be produced by expression

CC of PS118 polynucleotides in transfected host cells, and

CC immunogenic peptides are useful for raising PS118-specific

CC antibodies of diagnostic use. The methods and reagents of the

CC invention may provide an early means of detecting diseases of the

CC prostate and may also provide new markers which can differentiate

CC between the clinically important and unimportant prostate cancers

CC without the use of surgery.

XX Sequence 40 AA;

SO

Query Match 8.0%; Score 215; DB 23; Length 40;
Best Local Similarity 100.0%; Pred. No. 2.1e-13;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 332 SPKVEKDPSPRKKEMWENAGNKIYTMADKTIISKLTMEYK 371
DB 1 SPKVEKDPSPRKKEMWENAGNKIYTMADKTIISKLTMEYK 40

Search completed: July 24, 2003, 11:53:36
Job time : 170.369 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: July 24, 2003, 11:53:53 ; Search time 46.2635 Seconds
(without alignment)
473.743 Million cell updates/sec

Title: US-09-991-681-27

Perfect score: 2698
Sequence: 1 RIRAMQOVMLDTQCSPKT.....VRQAVREMLGRVGRVYDIIV 518

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA: *
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep: *
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep: *
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep: *
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep: *
5: /cgn2_6/ptodata/1/1aa/PTCUS.COMB.pep: *
6: /cgn2_6/ptodata/1/1aa/backflist.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2698	100.0	518	4	US-09-065-383-27
2	260	9.6	49	4	US-09-065-383-31
3	216	8.0	41	4	US-09-065-383-28
4	215	8.0	40	4	US-09-065-383-30
5	180	6.7	35	4	US-09-065-383-29
6	111	4.1	474	3	US-08-729-416C-1
7	111	4.1	474	3	US-09-433-353-1
8	111	4.1	594	3	US-08-729-416C-7
9	111	4.1	594	3	US-09-433-353-7
10	110	4.1	829	1	US-07-670-611-2
11	110	4.1	829	1	US-08-220-674-2
12	110	4.1	829	1	US-08-445-186-2
13	110	4.1	829	1	US-08-445-549-2
14	110	4.1	829	2	US-08-446-550-2
15	103	3.8	1098	2	US-08-923-992A-8
16	101.5	3.8	2101	1	US-08-466-390-4
17	101.5	3.8	2101	1	US-08-470-950-4
18	101.5	3.8	2101	1	US-08-467-781-4
19	101.5	3.8	2101	2	US-08-483-924-4
20	101.5	3.8	2101	3	US-09-452-294-1
21	100.5	3.7	2101	1	US-08-195-487-4
22	100.5	3.7	2101	5	PCT-US93-06160-4
23	98	3.6	693	4	US-09-252-991A-24059
24	98	3.6	733	3	US-08-725-459B-21
25	97.5	3.6	1128	3	US-08-922-992A-6
26	96	3.6	253	3	US-09-328-352-8134
27	95.5	3.5	1164	3	US-08-923-992A-2

28	95.5	3.5	1713	3	US-08-600-982-24	Sequence 24, Appl
29	95.5	3.5	1713	5	PCT-US94-10261A-24	Sequence 24, Appl
30	95.5	3.5	1805	1	US-07-853-913-2	Sequence 2, Appl
31	94.5	3.5	907	3	US-08-990-140-4	Sequence 4, Appl
32	94.5	3.5	907	4	US-09-546-238-4	Sequence 4, Appl
33	94.5	3.5	940	3	US-08-810-712-7	Sequence 7, Appl
34	94.5	3.5	1579	3	US-08-755-587-184	Sequence 184, App
35	94	3.5	397	4	US-09-006-428A-19	Sequence 2, Appl
36	94	3.5	2391	4	US-09-006-428A-19	Sequence 19, Appl
37	94	3.5	2391	3	US-08-466-855A-2	Sequence 2, Appl
38	94	3.5	2391	3	US-09-150-741-2	Sequence 2, Appl
39	93.5	3.5	984	1	US-08-257-073-3	Sequence 3, Appl
40	93.5	3.5	984	2	US-08-184-009-120	Sequence 120, App
41	93.5	3.5	984	2	US-08-458-356-120	Sequence 120, App
42	93.5	3.5	984	3	US-08-460-736-120	Sequence 120, App
43	93.5	3.5	984	4	US-09-535-370-120	Sequence 120, App
44	93.5	3.5	989	4	US-08-213-419B-2	Sequence 2, Appl
45	93.5	3.5	989	4	US-08-213-419B-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-09-065-383-27
Sequence 27, Application US/09065383
Patent No. 6391543
GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GORDON, JULIAN
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KLAAS, MICHAEL R.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: ROBERTS-RAVP, LISA
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
FOR DETECTING DISEASES OF THE PROSTATE
TITLE OF INVENTION: FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/065,383
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/842,385
FILING DATE: 23-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6084, US, P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 518 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6391543e
US-09-065-383-27

Query Match 100.0%; Score 2698; DB 4; Length 518;
Best Local Similarity 100.0%; Pred. No. 2.6e-277;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIRAMAQVFMIDTCCSPKTPNNFPAHSCQIILIEPPEKNGHTKKSVPREIIVSL 60
DB 1 RIRAMAQVFMIDTCCSPKTPNNFPAHSCQIILIEPPEKNGHTKKSVPREIIVSL 60
QY 61 SHQVLLQNLVYDILLBEFVKGPSGEEKTIQVPEAKIAGELRYISQNLAVIFDILLDSYR 120
DB 61 SHQVLLQNLVYDILLBEFVKGPSGEEKTIQVPEAKIAGELRYISQNLAVIFDILLDSYR 120
QY 121 TARFEDTSPGLKCLIKKYSIGIGANLYQASMSFNIFHALVCAVLNQETITAEQYK 180
DB 121 TARFEDTSPGLKCLIKKYSIGIGANLYQASMSFNIFHALVCAVLNQETITAEQYK 180
QY 181 VLFEDDERSTDSQCCSSEDEIFETAQVSPRGKEKQMPARPLSVQPVSNADWVW 240
DB 181 VLFEDDERSTDSQCCSSEDEIFETAQVSPRGKEKQMPARPLSVQPVSNADWVW 240
QY 241 LVKRLHLKLMELCNVYIQMHDLNENCMEEPPIFKGDPEFILPSFQSESSTPSTGPGSGKE 300
DB 241 LVKRLHLKLMELCNVYIQMHDLNENCMEEPPIFKGDPEFILPSFQSESSTPSTGPGSGKE 300
QY 301 TPSEBDRSOSREHMGESLSIKAGGDLILPSPKYEKDPSRKKEWENAGKITYTMAD 360
DB 301 TPSEBDRSOSREHMGESLSIKAGGDLILPSPKYEKDPSRKKEWENAGKITYTMAD 360
QY 361 KTISKLMEYKRRKQOHNLSAPFKEVKEKGEPLGPRGODSPLLORPQHLMDQGMHS 420
DB 361 KTISKLMEYKRRKQOHNLSAPFKEVKEKGEPLGPRGODSPLLORPQHLMDQGMHS 420
QY 421 FSAGPELLRQDKRPPSGSTGSSLSVSRDAEAOIQAMTMVLTINQIOLPDQFTALQ 480
DB 421 FSAGPELLRQDKRPPSGSTGSSLSVSRDAEAOIQAMTMVLTINQIOLPDQFTALQ 480
QY 481 PAVFPCISOLTCAMTDIRVQAVREMLGRVGRVYDIIV 518
DB 481 PAVFPCISOLTCAMTDIRVQAVREMLGRVGRVYDIIV 518

RESULT 2
US-09-065-383-31
Sequence 31, Application US/09065383
Patent No. 6391543

GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GORDON, JULIAN
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KLAS, MICHAEL R.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: ROBERTS-RAPP, LISA
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA

ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/065,383
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/842,385
FILING DATE: 23-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6084.US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6391543e
US-09-065-383-31

Query Match 9.6%; Score 260; DB 4; Length 49;
Best Local Similarity 100.0%; Pred. No. 2.9e-20;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 EPLPRGODSPLLORPQHLMDQGMHSFSAGPELLRQDKRPPSGSTGS 441
DB 1 EPLPRGODSPLLORPQHLMDQGMHSFSAGPELLRQDKRPPSGSTGS 49

RESULT 3
US-09-065-383-28
Sequence 28, Application US/09065383
Patent No. 6391543

GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GORDON, JULIAN
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KLAS, MICHAEL R.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: ROBERTS-RAPP, LISA
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/065,383

FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/842,385
FILING DATE: 23-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6084.US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 41 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6391543e
US-09-065-383-28

Query Match 8.0%; Score 216; DB 4; Length 41;
Best Local Similarity 100.0%; Pred. No. 1e-15;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 184 EDDERSTDSQCCSEDEDIFEEETAOVSPPRGKKEKROWRAR 224
Db 1 EDDERSTDSQCCSEDEDIFEEETAOVSPPRGKKEKROWRAR 41

RESULT 4

US-09-065-383-30
Sequence 30, Application US/09065383
Patent No. 6391543

GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GORDON, JULIAN
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KLAS, MICHAEL R.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: ROBERTS-RAPP, LISA
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
FILING DATE: US/09/065,383
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/842,385
FILING DATE: 23-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441

REFERENCE/DOCKET NUMBER: 6084.US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:

INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6391543e
US-09-065-383-30

Query Match 8.0%; Score 215; DB 4; Length 40;
Best Local Similarity 100.0%; Pred. No. 1.3e-15;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 332 SPVKEKKDPSRKKEWENAGNKIYTMAADKTSKLTETK 371
Db 1 SPVKEKKDPSRKKEWENAGNKIYTMAADKTSKLTETK 40

RESULT 5

US-09-065-383-29
Sequence 29, Application US/09065383
Patent No. 6391543

GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GORDON, JULIAN
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KLAS, MICHAEL R.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: ROBERTS-RAPP, LISA
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
FILING DATE: US/09/065,383
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/842,385
FILING DATE: 23-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6084.US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:

INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 35 amino acids

```

; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: NO. 6391543e
US-09-065-383-29

Query Match
Best Local Similarity 6.7%; Score 180; DB 4; Length 35;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 283 SFQSSSTPSTGFGSKETPSDDRSQSRHMGES 317
DB 1 SFQSSSTPSTGFGSKETPSDDRSQSRHMGES 35

RESULT 6
US-08-729-416C-1
; Sequence 1, Application US/08729416C
; Patent No. 6013767
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, TAKESHI
; TITLE OF INVENTION: NOVEL BRAIN-SPECIFIC ADAPTER MOLECULE GENE
; TITLE OF INVENTION: THEREOF, AND ANTIBODY THERETO
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PILLSBURY, MADISON & SUTRO, L.L.P.
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/729,416C
; FILING DATE: 11-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRY, GLENN J.
; REGISTRATION NUMBER: 28458
; REFERENCE/DOCKET NUMBER: 7898/225948
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 474 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-729-416C-1

Query Match
Best Local Similarity 4.1%; Score 111; DB 3; Length 474;
Matches 90; Conservative 71; Mismatches 205; Indels 120; Gaps 19;

QY 83 PGEKTIQVPE-----AKLAGFLRYI-SMQLAVIFDILLDSYRTARE 124
DB 9 PDEPLPRPRGTPHASDQVLGPGVTYVVKYLGCEIVLRSMRSLDFSTRQITREAIRSV 68
QY 125 FDTSPGLKCLKLK-----VSGIGGANL-YROSAMFNIFYHALVCAVLTNQETITAE 176
DB 69 CEAVGAGAKAFKRRKPPSKMLSSILGKSNLQFAGMSISLTISTASINLRTPDSKQIIANH 128
QY 177 QYKXVLFE-----DERSTDSSQCCSDEDEIFFEITAOVSPRGKEX 218
DB 129 HHRISIFASGGDPDTTDDVAVYAKDPVNRACHILLECCDGLAODVIGSIGAFELRFRQY 188
QY 219 RQWRARMLPLSVQPVSNADWVWLVRKLHKLQMLCNNTYIQMLDLENCEMEBPPIKGPDPF 278

```

```

DB 189 LQCPKIPALH-DMQSIDPEW-----TEEBGDGDHPYINSPS 227
QY 279 -----FILPSFQSSSTPSTGFGSKETPSDDRSQSRHMGESLS-----LXAGG 324
DB 228 KMPPGGFLDTRLRKPRHAPDPAQFAGEQYVYGR-----HLGDTGEBDWQOTPLRQGS 282
QY 325 GDLLPSPKYEKKDPSEKKEWENAGNKITYMAADTISKLMTEYKKRQOHNLSAPPK 384
DB 283 SDIVSTPEGKLHVA-PTGEAPTYVNT-QQIPQAMPAAVSS--AESSPKDLFDMKEPFD 338
QY 385 EVKYEKKG-----EPLGRGQDSPLQRPQ-HLMDGOMHSPSAGPELLROD 431
DB 339 ALKQOPLQPVLSKASVCECISPVSPRAPDAMLEELQAEITYQSEMSRKAEG--LLEKD 396
QY 432 ----KRPSSGSTGSLSVYDAEAOIQAWTNVLTLYNIOILPDQTFPALQPAVPCPI 487
DB 397 GDPLVRKSTNPGSFVLGMHNGQAK-----HLLVDPREGTIRT-KQVFDPSI 443
QY 488 SOLTCH 493
DB 444 SHLINH 449

RESULT 7
US-09-433-353-1
; Sequence 1, Application US/09433353
; Patent No. 6545141
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, TAKESHI
; TITLE OF INVENTION: NOVEL, BRAIN -SPECIFIC ADAPTER MOLECULE, GENE THEREOF,
; TITLE OF INVENTION: AND ANTIBODY THERETO
; FILE REFERENCE: 7898/262241
; CURRENT APPLICATION NUMBER: US/09/433,353
; PRIOR FILING DATE: 1999-11-04
; PRIOR APPLICATION NUMBER: 08/729,416
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 474
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-433-353-1

Query Match
Best Local Similarity 4.1%; Score 111; DB 4; Length 474;
Matches 90; Conservative 71; Mismatches 205; Indels 120; Gaps 19;

QY 83 PGEKTIQVPE-----AKLAGFLRYI-SMQLAVIFDILLDSYRTARE 124
DB 9 PDEPLPRPRGTPHASDQVLGPGVTYVVKYLGCEIVLRSMRSLDFSTRQITREAIRSV 68
QY 125 FDTSPGLKCLKLK-----VSGIGGANL-YROSAMFNIFYHALVCAVLTNQETITAE 176
DB 69 CEAVGAGAKAFKRRKPPSKMLSSILGKSNLQFAGMSISLTISTASINLRTPDSKQIIANH 128
QY 177 QYKXVLFE-----DERSTDSSQCCSDEDEIFFEITAOVSPRGKEX 218
DB 129 HHRISIFASGGDPDTTDDVAVYAKDPVNRACHILLECCDGLAODVIGSIGAFELRFRQY 188
QY 219 RQWRARMLPLSVQPVSNADWVWLVRKLHKLQMLCNNTYIQMLDLENCEMEBPPIKGPDPF 278
DB 189 LQCPKIPALH-DMQSIDPEW-----TEEBGDGDHPYINSPS 227
QY 279 -----FILPSFQSSSTPSTGFGSKETPSDDRSQSRHMGESLS-----LXAGG 324
DB 228 KMPPGGFLDTRLRKPRHAPDPAQFAGEQYVYGR-----HLGDTGEBDWQOTPLRQGS 282
QY 325 GDLLPSPKYEKKDPSEKKEWENAGNKITYMAADTISKLMTEYKKRQOHNLSAPPK 384
DB 283 SDIVSTPEGKLHVA-PTGEAPTYVNT-QQIPQAMPAAVSS--AESSPKDLFDMKEPFD 338

```



```

Db      459 ALKNQPLGVLTKASAVESICSPVSPRAPDAKYLEELQAEWTYQGMRSKREAG--LLEKD 516
Qy      432 ----KPPSGSTGSSLSVSRDAEAOIQAMTMVLTVLNOIQILPDQFTTALQPAVPCFI 487
Db      517 GDFLVKSTPNPGSFVLTCMHNGQAK-----HLILVDEGRTIRT-KDRVDFSI 563
Qy      488 SOLTCH 493
Db      564 SHLINH 569

RESULT 10
US-07-670-611-2
; Sequence 2, Application US/07670611
; Patent No. 5330892
; GENERAL INFORMATION:
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth W.
; APPLICANT: White, Raymond
; APPLICANT: Nakamura, Yusuke
; TITLE OF INVENTION: Gene Mutated in Colorectal Cancer of
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner, Birch et al.
; STREET: 1001 G Street
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20001-4597
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/670,611
; FILING DATE: 19910313
; CLASSIFICATION: 435.
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 1107.33981
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 829 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; US-07-670-611-2

Query Match      4.1%; Score 110; DB 1; Length 829;
Best Local Similarity 18.4%; Pred. No. 0.028;
Matches 92; Conservative 81; Mismatches 198; Indels 130; Gaps 19;
Qy      81 PEPGEKTI-----QVPEALAGFLRYISQNLAVIFDLILDSRYRAREDDTSFGKCLIK 136
Db      307 PGTGLSTSSSSNDPIPIAKIARVK-----LSKTRSSSSSDRPVIGS 349
Qy      137 KVSIGGAANTLYRQAMSF-----NIYFHALVCAVLTNQETITAEQKVLFPED 185
Db      350 EISSISGVSSVAEHLAHSIQDSNIOETIQTLYSHG-----SAISSEKIRE--FEV 398
Qy      186 DERSTDDSQCCSEDEDIFEEETAQVSPPRGKEKROWRAMPLLSVQPVSNADWVLVYRL 245
Db      399 FTERLNSRIEHLKSNQDLITLITL-----ECKSNAERMSMLVGKYESNATATLALQY 451

```

```

Qy      246 HKLMELCNNTYIQMHLENCMEBPPIFKGDPPIILPSFQSEESTPSTGCFSGKETPSE- 304
Db      452 SEQDIEAYELL-----ALAESEQSLLIQOFPAAGVSSPGQSGDENTTOM 498
Qy      305 ----DPRSQRHEHGESISLKAGS--GDLPLPSPKYEKDPSRKEMENAGKITYMA 358
Db      499 LKRAHDCRKTRENAKALLMKLKDSCCGAPAVAGCSVP-----WESLSNSHTST 549
Qy      359 ADKTIKLMTEYKRRKQOHNLSAPKEVYKKG-----EPIG---PRGQD 401
Db      550 TSSTASSCDTEFTKEDBQ-RLKDYIQQLKNDRAAVKLTMLESIHIDPLSYDVKPRGDS 608
Qy      402 SPL-LQRP---QHMDGQGRHSFSAGPELLQDKRPRSSGSTSSLSVSRDAEAOIQAM 457
Db      609 QRLDLENVAVIQEIMANKKEEMAEIKQYLLEKEK-----ALELKLSTREAQOAY 660
Qy      458 TNNVLTVINOIQILPDQFTTALQ-----PAVFPCIS---QLTCHVDIR 498
Db      661 LVHIEHLKSEVEEQEQRMSLSSTSSGSKDQPKGECADAPALSLAEIRTTCSENEELA 720
Qy      499 VR--QAVREWLGRVGRVYDII 517
Db      721 AEFTNAIRREKKLKAQVQELV 741

RESULT 11
US-08-220-674-2
; Sequence 2, Application US/08220674
; Patent No. 5571805
; GENERAL INFORMATION:
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth W.
; APPLICANT: White, Raymond
; APPLICANT: Nakamura, Yusuke
; TITLE OF INVENTION: Gene Mutated in Colorectal Cancer of
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner, Birch et al.
; STREET: 1001 G Street
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20001-4597
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/220,674
; FILING DATE: 31-MAR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/670,611
; FILING DATE: 13-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 1107.33981
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 829 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; ORIGINAL SOURCE:

```

ORGANISM: Homo sapiens
US-08-220-674-2

4.1%; Score 110; DB 1; Length 829;

Query Match
Best Local Similarity 18.4%; Pred. No. 0.028; Indels 130; Gaps 19;
Matches 92; Conservative 81; Mismatches 198;

```

QY 81 PPSPEEKTI-----QVPEAKLAGFLRYISMQNLAVIPDLILDYRTAREPDTSPGLKCLK 136
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 307 PSTGLSTSSSSNDIPAKIAERK-----LSKTSSESSSDRPVLGS 349
QY 137 KVSIGIGANILYROSAMSF-----NIYFHALVCAVLTNOETITAEQVKVLFED 185
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 350 EISSIGVSSVAHLNLSLDDCSNIOEIFQTLVSHG-----SAISSEKIRE--FEV 398
QY 186 DERSTDSQCCSDEDEIFEEITAVSPRGKEKQWRARMPLLSVQPSNADWVWLKRL 245
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 399 ETERLNSRIEHLKSONDLITLLE-----ECKSNMERMSMLVGKYESNATRLRLAQ 451
QY 246 HKLCMELCNNTYQMHLDLNCMEEPPIFKGDPFPIPSFQSESSSTPTGSGKETPSE- 304
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 452 SEQCIENAYELL-----ALAESQSLILGQFRAAGVSSFGDQSGDENITQM 498
QY 305 ----DERSQSRHNGESISLAKG--GDLLPSPPKVEKKDPSRKKEWENAGKIYMA 358
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 499 LKRAHDCRKTAEENAKALMLKLDSCGAFVAGCSVQP-----WESLSNSHTST 549
QY 359 ADKTIKLTMEYKRRKQOHNLSAPPKVEKKG-----EPLG----PRGQD 401
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 550 TSSYASSCDTEFTKDEQ--RLKDYIQOLKNDRAAVKLTMLELSIHIDPLSYDVKPRGDS 608
QY 402 SPL-LQRP---QHLMDOGQMRHSFSGAPPELLRODKRPRSGSTGSSLSVSRDAEAQIQAW 457
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 609 QRLDENAVLMQELMAMKEEMAEIKAQYILLEKEK-----ALELKSTRDAQOAY 660
QY 458 TNNVLTVALNQIQLPDQTFALQ-----PAVPCIS-----QLTCHVTDIR 498
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 661 LVHIEHLKSEVEBEKQRMRSLSSTSGSKDKPKECADAAPALSIAELRTTCSSEMLA 720
QY 499 VR--QAVREWLGRVGRVYDII 517
Db 721 AEFTHAIRREKKLKARVOELV 741

```

RESULT 12

US-08-445-186-2
Sequence 2, Application US/08445186

Patent No. 5576422
GENERAL INFORMATION:
APPLICANT: Vogelstein, Bert
APPLICANT: Kinzler, Kenneth W.
APPLICANT: White, Raymond
APPLICANT: Nakamura, Yusuke
TITLE OF INVENTION: Gene Mutated in Colorectal Cancer of
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner, Birch et al.
STREET: 1001 G Street
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20001-4597
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/445.186
FILING DATE: 19-MAY-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/220,674
FILING DATE: 31-MAR-1994
APPLICATION NUMBER: US 07/670,611
FILING DATE: 13-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 1107.33981
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 829 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: YES
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-08-445-186-2

Query Match
Best Local Similarity 18.4%; Pred. No. 0.028; Indels 130; Gaps 19;
Matches 92; Conservative 81; Mismatches 198;

```

QY 81 PPSPEEKTI-----QVPEAKLAGFLRYISMQNLAVIPDLILDYRTAREPDTSPGLKCLK 136
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 307 PSTGLSTSSSSNDIPAKIAERK-----LSKTSSESSSDRPVLGS 349
QY 137 KVSIGIGANILYROSAMSF-----NIYFHALVCAVLTNOETITAEQVKVLFED 185
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 350 EISSIGVSSVAHLNLSLDDCSNIOEIFQTLVSHG-----SAISSEKIRE--FEV 398
QY 186 DERSTDSQCCSDEDEIFEEITAVSPRGKEKQWRARMPLLSVQPSNADWVWLKRL 245
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 399 ETERLNSRIEHLKSONDLITLLE-----ECKSNMERMSMLVGKYESNATRLRLAQ 451
QY 246 HKLCMELCNNTYQMHLDLNCMEEPPIFKGDPFPIPSFQSESSSTPTGSGKETPSE- 304
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 452 SEQCIENAYELL-----ALAESQSLILGQFRAAGVSSFGDQSGDENITQM 498
QY 305 ----DERSQSRHNGESISLAKG--GDLLPSPPKVEKKDPSRKKEWENAGKIYMA 358
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 499 LKRAHDCRKTAEENAKALMLKLDSCGAFVAGCSVQP-----WESLSNSHTST 549
QY 359 ADKTIKLTMEYKRRKQOHNLSAPPKVEKKG-----EPLG----PRGQD 401
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 550 TSSYASSCDTEFTKDEQ--RLKDYIQOLKNDRAAVKLTMLELSIHIDPLSYDVKPRGDS 608
QY 402 SPL-LQRP---QHLMDOGQMRHSFSGAPPELLRODKRPRSGSTGSSLSVSRDAEAQIQAW 457
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 609 QRLDENAVLMQELMAMKEEMAEIKAQYILLEKEK-----ALELKSTRDAQOAY 660
QY 458 TNNVLTVALNQIQLPDQTFALQ-----PAVPCIS-----QLTCHVTDIR 498
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 661 LVHIEHLKSEVEBEKQRMRSLSSTSGSKDKPKECADAAPALSIAELRTTCSSEMLA 720
QY 499 VR--QAVREWLGRVGRVYDII 517
Db 721 AEFTHAIRREKKLKARVOELV 741

```

RESULT 13

US-08-446-549-2
Sequence 2, Application US/08446549

Patent No. 5693536
GENERAL INFORMATION:
APPLICANT: Vogelstein, Bert
APPLICANT: Kinzler, Kenneth W.
APPLICANT: White, Raymond
APPLICANT: Nakamura, Yusuke

```

/ TITLE OF INVENTION: Gene Mutated in Colorectal Cancer of
/ TITLE OF INVENTION: Humans
/ NUMBER OF SEQUENCES: 19
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Banner, Birch et al.
/ STREET: 1001 G Street
/ CITY: Washington
/ STATE: D.C.
/ COUNTRY: U.S.A.
/ ZIP: 20001-4597
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/446,549
/ FILING DATE: 19-MAY-1995
/ CLASSIFICATION: 514
/ PRIORITY APPLICATION DATA:
/ APPLICATION NUMBER: US 07/670,611
/ FILING DATE: 13-MAR-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Kagan, Sarah A.
/ REGISTRATION NUMBER: 32,141
/ REFERENCE/DOCKET NUMBER: 1107.33981
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 202-508-9100
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 829 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHEICAL: YES
/ ANTI-SENSE: NO
/ ORIGINAL SOURCE:
/ ORGANISM: Homo sapiens
/ US-08-446-549-2

Query Match      4.1%; Score 110; DB 1; Length 829;
Best Local Similarity 18.4%; Pred. No. 0.028;
Matches 92; Conservative 81; Mismatches 198; Indels 130; Gaps 19;

QY 81 PSPGEKTI---OVPEAKLAGFLRYISMQNLAVIFDILLDSYRTAREFTSPGKCLK 136
DB 307 PSTGEISTSSSNDIPAKIAERVK-----LSKTRSSSSSDRPVLGS 349
QY 137 KVSIGGANLYRQSMF-----NIYFHALVCAVLTNQETITAEQVKVLPED 185
DB 350 EISSIGVSSVAEHLAHSLODCSNIQEIFQTLVSHG-----SAISESKIRE--FEV 398
QY 186 DERSTDSQQCSSEDEDIPEETAQVSPPRGKEKQWARMPLLSVQPVSNADWWLVKRL 245
DB 399 ETERLNSRIEHLKSNQDLITLLE-----ECKSAEMSMVLGKYBESNATALLALOY 451
QY 246 HKLCMELCNVYIQMLDLLENCEMEEPPIFKGDPFLLPSFQSSSTPSTGFSKGPSPR- 304
DB 452 SEQCTEAYELL-----ALASEQSLILGQPRAGVSSPEDQSGDENITQM 498
QY 305 ---DDRSGSRHEMGSSLSLKAGG--GDLILPPSPVEKKDPSRKKEWENAKKIYMA 358
DB 499 LKRAHDCRTAENNAKALMLKLDGSCGAFVAGCVDP-----WESLSNSHST 549
QY 359 ADKTSIKMTEYKKRQOHNLSAFPKEVVEKKG-----EPLG---PRGD 401
DB 550 TSTSTASCTEFTKEBEO-RLKDYIQQLKNDRAAVLTLLESLIHIDLSIDVKEPRGDS 608
QY 402 SPL-LQRP---QHLMDQGMRRHSFSNAPPELLRODKRPRSGSTGSSLSVSDRAAQIQAM 457
DB 609 QRLDLENNAVLMQELMAMKEEMALIKAQIYLKEKK-----ALTELKSTRDAQEOQVY 660

```

```

QY 458 TMMVLTVLNOQIILPDQTFRLQ-----PAVFPCIS---QLTCHVTDIR 498
DB 661 LVHIEHLKSEVEEOKRMRLSTSSGSKDKPGKCADAPSLSLAEIRTTCSNELA 720
QY 499 VR--QAVREMLGRVGRVYDII 517
DB 721 AEFTHAIRREKKRLARVOELV 741

RESULT 14
US-08-446-550-2
/ Sequence 2, Application US/08446550
/ Patent No. 5830676
/ GENERAL INFORMATION:
/ APPLICANT: Vogelstein, Bert
/ APPLICANT: Kinzler, Kenneth W.
/ APPLICANT: White, Raymond
/ APPLICANT: Nakamura, Yusuke
/ TITLE OF INVENTION: Gene Mutated in Colorectal Cancer of
/ NUMBER OF SEQUENCES: 19
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Banner, Birch et al.
/ STREET: 1001 G Street
/ CITY: Washington
/ STATE: D.C.
/ COUNTRY: U.S.A.
/ ZIP: 20001-4597
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/446,550
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIORITY APPLICATION DATA:
/ APPLICATION NUMBER: US/07/670,611
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Kagan, Sarah A.
/ REGISTRATION NUMBER: 32,141
/ REFERENCE/DOCKET NUMBER: 1107.33981
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 202-508-9100
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 829 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHEICAL: YES
/ ANTI-SENSE: NO
/ ORIGINAL SOURCE:
/ ORGANISM: Homo sapiens
/ US-08-446-550-2

Query Match      4.1%; Score 110; DB 2; Length 829;
Best Local Similarity 18.4%; Pred. No. 0.028;
Matches 92; Conservative 81; Mismatches 198; Indels 130; Gaps 19;

QY 81 PSPGEKTI---OVPEAKLAGFLRYISMQNLAVIFDILLDSYRTAREFTSPGKCLK 136
DB 307 PSTGEISTSSSNDIPAKIAERVK-----LSKTRSSSSSDRPVLGS 349
QY 137 KVSIGGANLYRQSMF-----NIYFHALVCAVLTNQETITAEQVKVLPED 185
DB 350 EISSIGVSSVAEHLAHSLODCSNIQEIFQTLVSHG-----SAISESKIRE--FEV 398
QY 186 DERSTDSQQCSSEDEDIPEETAQVSPPRGKEKQWARMPLLSVQPVSNADWWLVKRL 245

```



```

Db      399 ETERLINSRIEHLKQNDLTLTLR-----ECKSNMRSMVLGYKESNATLRLALQY 451
QY      246 HKLCELCNNYIQHNLLENCEMEEPPIFKGDPFPLPSFOSESSTPSTGSGKETPSE- 304
Db      452 SEQCIIEAVELL-----ALAESQSLILGQFRAGYSSGSDGSDGENTTOM 498
QY      305 ----DBRSQSNHNGESLSLKAG--GDLLPSPSPKYEKDPSPKKEWENAGNKIYTMA 358
Db      499 LKRAHDKRKTENAKALLMLKLDSCGGAFAVACGSVQ-----WELSSNSHTST 549
QY      359 ADKTIKSLMTEYKQRKQOHLSAPPEYKVEKKG-----EPLG----PRGQD 401
Db      550 TSSSTASCDTFTFEDEQ-RLKDYIQQLKNDRAAVKLTMLEIESIHIDPLSYDVKPRGDS 608
QY      402 SPL-LORP---OHLMDGOMHSPSAGPELLRQDKRPRSSGSSLSVSVDAEAQIQAW 457
Db      609 ORLDENAVLMQELMAMKEEAEKALQYLLEKEK-----ALEIKSTREAQEQAY 660
QY      458 TNNVLTVALNOIQLPDQTFALQ-----PAVFPGIS-----QLTCHVTDIR 498
Db      661 LVHIEHLKSEVEEQEQRMRSLSTSSGSKDKPGKECADNASPALSLAELRTTSENELA 720
QY      499 VR--QAVREMLGRGRVDTI 517
Db      721 AEFTHAIRREKKLKAQVQELV 741

```

RESULT 15

```

US-08-923-992A-8
; Sequence 8, Application US/08923992A
; Patent No. 6280738
; GENERAL INFORMATION:
; APPLICANT: Tai, Joseph Y.
; TITLE OF INVENTION: No. 6280738-IGA Fc Binding Forms of the Group B
; TITLE OF INVENTION: Streptococcal Beta Antigens
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/923,992A
; FILING DATE: 05-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024,707
; FILING DATE: 06-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bemond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 1438.0140001/RWE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1098 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-923-992A-8

```

Query Match 3.8%; Score 103; DB 3; Length 1098;
 Best Local Similarity 23.7%; Pred. No. 0.25;

```

Matches 103; Conservative 66; Mismatches 177; Indels 88; Gaps 22;
QY      38 PDEKPNHTKKSVPFELIVSLSHQVLLQNLVDILLEEVKGPSPGEEKTIOVPEAKLA 97
Db      199 PDKKEDAEV-----VBEELKTLFSS---TGAIDQEIQEHVKKETSSSEENTQKDE---- 247
QY      98 GFLRYI-SMONTAVIFDLIDSYRTAREPDTSPGLKCL-----LKVSGIGGANLYROS 151
Db      248 ---HYANSLQNLAKQSLBEIDK-ATTNEQATQYKNGQLBNAQQLKEIPLIKETNVLYK 303
QY      152 AMSFNIFYHALVCAVLTNQETTAEOYKVLFFEDDERSTD----SSQCCSSEDEDLFEET 207
Db      304 AMSESL-----EQYKELKHNSSEANLEDLVAKSKEIIVREYEGKLNQS 345
QY      208 AQVSPRGEKQRQWRAMPPLSV-----QPVSNADWVWLKRLHKLCELCNNYIQHNL 262
Db      346 KNL--PELKOLEE-EAHSKLLQYVEDPRKFKTSBOVTPKKRLKRLDLANENN--QOKIE 400
QY      263 LENCMEEPPIFKGDPFPLPSFOSESSTPSTGSPG---KETPSEDDR-----SOSREH 313
Db      401 LVTSPEKITVYEGEDVYFTVAKSDSKT--TLDFSDLTKYNPVSVDRISTNYKTNTDNH 458
QY      314 MGSLSLKAGGDLPLPSPKVEKQDPSRKKEWENAGNKIYTMAADKTIKLMTEYKQR 373
Db      459 KIAIITIK---NLKNESQTVTLKAK-----DSGNVY-----EKTFT--ITVQKKE 500
QY      374 KOQHNLSAPPEYKVEKKEGRLGPRGQDSPLOPQHLMD--QGOMHSPSAGPELLROD 431
Db      501 EKQVPTPEQKDSKTEBK-VQEPKSNQKQL---QELINSQOOLEKELKAIKELMEQEP 556
QY      432 KRPRSGSTGSSLSV 445
Db      557 EIPSNPEYGIQKSI 570

```

Search completed: July 24, 2003, 12:16:20
 Job time : 48.2635 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 11:57:13 ; Search time 59.9151 Seconds
(without alignments)
1026.746 Million cell updates/sec

Title: US-09-991-681-27

Perfect score: 2698
Sequence: 1 RIRAAQVFMVLTQCSPKT.....VRAVREMLGRVGRVYDIIV 518

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PC7_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/PC7US_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	110	4.1	828	US-08-681-219-28	Sequence 28, Appl
2	108.5	4.0	2783	US-09-816-669A-14	Sequence 14, Appl
3	102.5	3.8	453	US-10-106-698-5604	Sequence 5604, Ap
4	100.5	3.7	1294	US-09-836-499-2	Sequence 2, Appl
5	100.5	3.7	1294	US-10-162-435-2	Sequence 2, Appl
6	100.5	3.7	1309	US-09-836-499-5	Sequence 5, Appl
7	100.5	3.7	1309	US-10-162-435-5	Sequence 5, Appl
8	100	3.7	557	US-10-106-698-5295	Sequence 5295, Ap
9	98	3.6	411	US-09-788-600-5	Sequence 5, Appl
10	98	3.6	446	US-09-738-626-4700	Sequence 4700, Ap
11	98	3.6	1332	US-09-982-091A-4	Sequence 4, Appl
12	97.5	3.6	669	US-09-823-187-88	Sequence 88, Appl
13	96.5	3.6	759	US-09-764-864-823	Sequence 823, App
14	96	3.6	555	US-09-764-864-1161	Sequence 1161, Ap
15	96	3.6	555	US-09-764-864-1574	Sequence 1574, Ap

Result No.	Score	Query Match	Length	ID	Description
16	95.5	3.5	693	US-10-151-569-2	Sequence 2, Appl
17	95.5	3.5	1713	US-10-171-311-113	Sequence 113, Appl
18	95	3.5	678	US-09-823-187-87	Sequence 87, Appl
19	95	3.5	1395	US-10-153-668-468	Sequence 468, Appl
20	94.5	3.5	465	US-10-103-313-357	Sequence 357, Appl
21	94.5	3.5	504	US-09-801-368-208	Sequence 208, Appl
22	94.5	3.5	561	US-10-156-761-14518	Sequence 14518, A
23	94.5	3.5	600	US-09-764-864-1282	Sequence 1282, Ap
24	94.5	3.5	907	US-09-954-043-4	Sequence 4, Appl
25	94.5	3.5	907	US-10-200-154-4	Sequence 4, Appl
26	94.5	3.5	940	US-10-102-806-678	Sequence 678, Appl
27	94.5	3.5	1711	US-09-771-161A-219	Sequence 219, Appl
28	94.5	3.5	1711	US-09-771-161A-220	Sequence 220, Appl
29	94	3.5	258	US-09-815-242-5680	Sequence 5680, Ap
30	94	3.5	300	US-09-815-242-12173	Sequence 12173, A
31	93.5	3.5	304	US-09-987-107-7	Sequence 7, Appl
32	93.5	3.5	323	US-09-987-107-58	Sequence 58, Appl
33	93.5	3.5	519	US-09-934-455-164	Sequence 164, Appl
34	93	3.4	266	US-09-864-761-37015	Sequence 37015, A
35	93	3.4	464	US-09-902-941-1934	Sequence 1934, Ap
36	93	3.4	464	US-10-017-754-1934	Sequence 1934, Ap
37	93	3.4	615	US-09-925-301-1094	Sequence 1094, Ap
38	93	3.4	693	US-10-029-217A-4	Sequence 4, Appl
39	93	3.4	705	US-10-154-386-2	Sequence 2, Appl
40	93	3.4	1038	US-09-908-500A-2	Sequence 2, Appl
41	93	3.4	4019	US-09-738-973-425	Sequence 425, App
42	93	3.4	4019	US-09-854-133-425	Sequence 425, App
43	93	3.4	4019	US-10-144-649A-445	Sequence 425, App
44	92.5	3.4	393	US-09-745-763-19	Sequence 19, Appl
45	92.5	3.4	707	US-10-225-486-57	Sequence 57, Appl

ALIGNMENTS

RESULT 1
US-08-681-219-28
Sequence 28, Application US/08681219
Publication No. US20020058607A1

GENERAL INFORMATION:
APPLICANT: Takaki Sato and Junn Yanagisawa
TITLE OF INVENTION: COMPOUNDS THAT INHIBIT THE INTERACTION BETWEEN
TITLE OF INVENTION: SIGNAL-TRANSDUCING PROTEINS AND THE GTP
TITLE OF INVENTION: (IPDZ/DHR) DOMAIN AND USES THEREOF
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/681,219
FILING DATE: 22-JUL-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/48962/JPM/JRM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 391-0525
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 828 amino acids
TYPE: amino acid
STRANDEDNESS: single

```

QY 82 SPEEKTIOVPEAKLAGELFARYISMONLAVITFDLLDSYRRAREBDPSGKCLLKKVSGI 141
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 6 SPEFRIAMITIPSELEJDSILKYSIDLQNTLA-----KSLGRANLRATKLL 48
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 142 GGAANLYROSASFNIYFHALVCALVTNQETITAEQVKVLFEEDERSTDSSQCCSSEDE 201
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 49 KALKGTYTKEAKGN-----ENQDSQTSASSCDETEI 81
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 202 DFEETAAQVSPRGKEKIQWRAPMELLSVCP-----VSNADVWLKYRLHLKLCME 251
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 82 QISNQGEARQDLGHVYTK-RRRCTYAVVDDSOQNHSEIKISNP-----TE 127
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 252 LCNVNYIQWHLDELNCEBPPIFKGDPEFILDPSFQSESTPSTGQFSKETPSEDRSQSR 311

```

```

Db      128 PGNHEKQESQDLRATKVP-----SP-----PDEHQEAMNAVSSGNRDSKYVSEKKSILYT 178
Qy      312 EHMGESLTKAGGDDLLPSPKVEKKDPSRKKEWMEAGKITYMADKTIISLMTMEYK 371
Db      179 DE-----SSKPGKPKRTAITTPNFKQHEAHFKEM-----ESIDQYIERKK 219
Qy      372 KRKOQHNLSAPPEKVEKKKG--EPLGPRGQ-----DSPLLQRPQHLMDQGMHRSFSNAP 425
Db      220 KHFEHNSMNLKQOPINKGQVTRTPVPRGRSLVASTPISQR-----RSQGR-----SCGP 270
Qy      426 E-----LLRODKRP--RSGSTGSSLSVSDAE 451
Db      271 ASQSTLGKSLKRSATSAKTGVRFSPATKDNE 304

RESULT 4
US-09-836-499-2
; Sequence 2, Application US/09836499
; Publication No. US20030027316M1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 16051A AND 16051B, NOVEL HUMAN PDZ
; FILE REFERENCE: 10448-043001
; CURRENT APPLICATION NUMBER: US/09/836,499
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: US 60/197,507
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1294
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-836-499-2

Query Match      3.7%; Score 100.5; DB 11; Length 1294;
Best Local Similarity 21.8%; Pred. No. 6.2;
Matches 111; Conservative 66; Mismatches 194; Indels 139; Gaps 24;

Qy      10 FMDLQCSPTKPNPFDAQSCQILIELP-----PDEKNGHTKKSVSFREIVV----- 57
Db      839 FNMVVRMTQNSPDI-----ELIISQSKGVGNNPDEKNGTANSGVSTDLISFGYQG 892
Qy      58 SLISH-----QVLLQNTYDILLLEFPVKG-----PSGGEKTIQVEPAK 95
Db      893 SLISHTDQDQDNTTEELDMAGVQSLVPRRLHQLSLPLKAGASSCPSPPEISAGEIYFVE 952
Qy      96 LA---GFLRY-----ISMQNLAVIFDILLDSYRTAREFPTSQKLCLKKVSIGIGAA 145
Db      953 LVKEDGTIGFSVTGINTSVYGGIYVKSIVPGCPAKESGQILQGRLL--QVDSG----- 1006
Qy      146 NLVYQASANSFNIYFHALVCAVLNTOETITA---EQVKVLFEDEDERSTSSQOCSSEDE 201
Db      1007 -----ILCG--LTHQAVQCLKGPQVRLVLE--RNVPRSTQCCPSAND 1047
Qy      202 DIFETIAQVSPPRGKEKQWAPMLSVQF-----VSNADWVLVYRLHKLCHMELCNN 255
Db      1048 SMGDERTAVS-----LVYALPGRBSSCVSYTDGPKFEVKLKKANGLGFS 1092
Qy      256 YIQHNLLENCEMEPPIFKGDPPFILPFSQESSSTPTSGFSGETPESEDDRSQREHMG 315
Db      1093 FVQM--EKESCSH-----LKSDLVRIKRLFPQCPAENKAIAGDIILAVNGRSTREGLIFQ 1146
Qy      316 ESLSLKAGGD---LLPSPKVEKKDPSRKKEW--ENAGNKITYMA--ADTKISLKM 367
Db      1147 EVLHLIRAPQPEVTLILCRPPGAL---PEMEQEWQTEBELSADKEFTTRATCTDSCSPIL 1203
Qy      368 TEYKKRQKQHNLSAPPEKVEKKGEPLGPRGQSPILLQRPQHLMDQGMK----- 418
Db      1204 DQEDSWRD---SASP-----DAGEGILGRPESS---QKAIREQWQONREPRWASSLT 1250

```

```

Qy      419 HSFSAGEPLLRODKRPRSGSTGSSLSVSVR 448
Db      1251 HSPESHPLCKLHQERDESTLATSLEKDVR 1280

RESULT 5
US-10-162-435-2
; Sequence 2, Application US/10162435
; Publication No. US20030096305A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; APPLICANT: Gluckemann, Maria Alexandra
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Bandaru, Rajasekhar
; APPLICANT: Leiby, Kevin R.
; TITLE OF INVENTION: NOVEL HUMAN MEMBRANE-ASSOCIATED PROTEIN AND
; FILE REFERENCE: 10448-189001
; CURRENT APPLICATION NUMBER: US/10/162,435
; PRIOR FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: US 09/836,499
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: PCT/US01/12420
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: US 60/197,507
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/891,008
; PRIOR FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: PCT/US01/19963
; PRIOR FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: US 60/214,220
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 09/860,868
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: PCT/US01/16013
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,674
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/886,429
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: PCT/US01/20055
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/213,963
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 10/041,406
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: PCT/US02/00275
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/260,286
; PRIOR FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: US 09/934,268
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/41811
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,612
; PRIOR FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1294
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-162-435-2

Query Match      3.7%; Score 100.5; DB 15; Length 1294;
Best Local Similarity 21.8%; Pred. No. 6.2;
Matches 111; Conservative 66; Mismatches 194; Indels 139; Gaps 24;

Qy      10 FMDLQCSPTKPNPFDAQSCQILIELP-----PDEKNGHTKKSVSFREIVV----- 57
Db      839 FNMVVRMTQNSPDI-----ELIISQSKGVGNNPDEKNGTANSGVSTDLISFGYQG 892

```

```

QY 58 SLTSH-----OVLQNLVYLILLEEFGV-----PSRGEKTIQVPAK 95
DB 893 SLTSHQODRNTTEELDMAGVSLVRLRHQSLFPLKAGAGSSCPSPISAGSIYVE 952
QY 96 LA---GFLRY-----ISMQNLAVIFDLIDSYRTAREFDTSPKJCLKKVSGIGAA 145
DB 953 LKKEGTGTFSTGTGINTSVPGGIYKVSIVPGGNAEKGQILQGDRL--QVDEV----- 1006
QY 146 NLVROSAMSFNIYFHALVCAVLTNOETITA---EQVKVLFEDDERSTDSQCSSDE 201
DB 1007 -----ILCG-LTHKQAVOCLKGPQVAVLVE--RRVPRSTQCCPSAND 1047
QY 202 DIFETIAQVPPRKEKQWRAMPILLSVQ-----VSNADWVLVRLHLCMELCNN 255
DB 1048 SMGDERTAVS-----LVTLALPGRPSSCVSVTDGPKFVKLKNANGIGFS 1092
QY 256 YIQMHLDENCMEEPPIFKGDPPFILPFSQESSPTSGFSGKETPSEDDRSQSRHMG 315
DB 1093 FVQM--EKESCSH-----LKSDLVRIKRLFPQPAENGAIAAGDIIILVNGRSTEGILFQ 1146
QY 316 ESLSLKAGGD---LLLPSPRYEKKDPSRKKEW--ENAGNKIYTMA--ADKTIKLM 367
DB 1147 EVLHLRGAPQEVTLTLCRPPGAL---PEMEQEWQTELSADKEFTFRTATCTDCTSPIL 1203
QY 368 TEYKRRKQOHNLSAPPEKVEKGEPLGPRGODSPILQRPQHLMDQGM----- 418
DB 1204 DQEDSWRD---SASP-----DAGEGIGLAPRESS--QKAIREAQWQNRERPWASSILT 1250
QY 419 HSFSAPELLRODKRPRSGTSGSSLSVSR 448
DB 1251 HSPESHPLCTLHQERDESTLATSLEKQVR 1280

```

```

RESULT 6
US-09-836-499-5
; Sequence 5, Application US/09836499
; Publication No. US20030027316A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 16051A AND 16051B, NOVEL HUMAN PDZ
; FILE REFERENCE: 10448-043001
; CURRENT APPLICATION NUMBER: US/09/836,499
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: US 60/197,507
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1309
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-836-499-5

```

```

Query Match 3.7%; Score 100.5; DB 11; Length 1309;
Best Local Similarity 21.8%; Pred. No. 6.4; Indels 139; Gaps 24;
Matches 111; Conservative 66; Mismatches 194;

```

```

QY 10 FMLDTQCSPEKTPNNPDHAGSCQILIELP-----PDEKNGHTKKSVEFRIIV----- 57
DB 839 FMMAVRMTQNSPDNI-----ELIISQSGVGGNNPDEKNGHTANSQVSTILSFGVQ 892
QY 58 SLTSH-----OVLQNLVYLILLEEFGV-----PSRGEKTIQVPAK 95
DB 893 SLTSHQODRNTTEELDMAGVSLVRLRHQSLFPLKAGAGSSCPSPISAGSIYVE 952
QY 96 LA---GFLRY-----ISMQNLAVIFDLIDSYRTAREFDTSPKJCLKKVSGIGAA 145
DB 953 LKKEGTGTFSTGTGINTSVPGGIYKVSIVPGGNAEKGQILQGDRL--QVDEV----- 1006
QY 146 NLVROSAMSFNIYFHALVCAVLTNOETITA---EQVKVLFEDDERSTDSQCSSDE 201
DB 1007 -----ILCG-LTHKQAVOCLKGPQVAVLVE--RRVPRSTQCCPSAND 1047

```

```

QY 202 DIFETIAQVPPRKEKQWRAMPILLSVQ-----VSNADWVLVRLHLCMELCNN 255
DB 1048 SMGDERTAVS-----LVTLALPGRPSSCVSVTDGPKFVKLKNANGIGFS 1092
QY 256 YIQMHLDENCMEEPPIFKGDPPFILPFSQESSPTSGFSGKETPSEDDRSQSRHMG 315
DB 1093 FVQM--EKESCSH-----LKSDLVRIKRLFPQPAENGAIAAGDIIILVNGRSTEGILFQ 1146
QY 316 ESLSLKAGGD---LLLPSPRYEKKDPSRKKEW--ENAGNKIYTMA--ADKTIKLM 367
DB 1147 EVLHLRGAPQEVTLTLCRPPGAL---PEMEQEWQTELSADKEFTFRTATCTDCTSPIL 1203
QY 368 TEYKRRKQOHNLSAPPEKVEKGEPLGPRGODSPILQRPQHLMDQGM----- 418
DB 1204 DQEDSWRD---SASP-----DAGEGIGLAPRESS--QKAIREAQWQNRERPWASSILT 1250
QY 419 HSFSAPELLRODKRPRSGTSGSSLSVSR 448
DB 1251 HSPESHPLCTLHQERDESTLATSLEKQVR 1280

```

```

RESULT 7
US-10-162-435-5
; Sequence 5, Application US/10162435
; Publication No. US20030096305A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; APPLICANT: Glucksmann, Maria Alexandra
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Kapeller-Liebermann, Rosana
; APPLICANT: Bandaru, Rajsekhar
; APPLICANT: Leiby, Kevin R.
; TITLE OF INVENTION: NOVEL HUMAN MEMBRANE-ASSOCIATED PROTEIN AND
; FILE REFERENCE: 10448-189001
; CURRENT APPLICATION NUMBER: US/10/162,435
; PRIOR FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: US 09/836,499
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: PCT/US01/12420
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: US 60/197,507
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/891,008
; PRIOR FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: PCT/US01/19963
; PRIOR FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: US 60/214,220
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 09/860,868
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: PCT/US01/16013
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,674
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/886,429
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: PCT/US01/20055
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/213,963
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 10/041,406
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: PCT/US02/00275
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/260,286
; PRIOR FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: US 09/934,268
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/41811
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,612

```

;; PRIOR FILING DATE: 2000-08-21
;; NUMBER OF SEQ ID NOS: 38
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 5
;; LENGTH: 1309
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-162-435-5

Query Match 3.7%; Score 100.5; DB 15; Length 1309;
Best Local Similarity 21.8%; Pred. No. 6.4;
Matches 111; Conservative 66; Mismatches 194; Indels 139; Gaps 24;

QY 10 FMILDTQCSPTKPTNPFDAQSCOLIIELEP-----PDEKPNHTKKSVSFREIVV----- 57
DB 639 FMAVMRMIONSPDNI-----ELIISQKGVGKGNPDDEKNGKTANSQVSDILSPFGQG 892
QY 58 SLASH-----QVLLQNLVDILLEEFVKG-----PSPGEKTIQVPEAK 95
DB 893 SLSTHDPDQDRNTEELDMAGVQSLVPRLRHQLSLFLPLKAGASSCPSPPEISAGEIYFVE 952
QY 96 LA---GFLRY-----ISMQNLAVIPDLLDVRYTAEPFTSGIKCLKKKVSGIGAA 145
DB 953 LVKEDGTLGFSVTGINTSVYGGIYVKSIVPGGPAKEGQILQGDRL--QVDGV--- 1006
QY 146 NLVRSAMSFNIYFHALVCAVLTNQETITA---EQVKVLFEDDERSTDSSQCCSSEDE 201
DB 1007 -----ILCG-LTHKQAVQCLKGPGQVAVLVE--RRVPRSTQCCPSAND 1047
QY 202 DIFETNAVSPPRKEKQWRAMPPLSVQ-----VSNADWWLVYKRLHKLCEMLCNN 255
DB 1048 SMGDERFVVS-----LVYALPGRPSSCVSVTDGPFVYKLKKNANGLOFS 1092
QY 256 YIQHMLDLENCEMEBPPIFKGDPFILPFSQSESTPSTGSGKETBEDRSQREHMG 315
DB 1093 FVQW--EKESCSH---LKSDLVRIKRLFGQPAEENCAIAGIITLAVNGRSTEGILFQ 1146
QY 316 ESLSKAGGAD---LLPSPKVEKKDPSRKKEMW--ENAGNKITYWA--ADTKISKLM 367
DB 1147 EVLHLRGAPQEVVLLLCRPPGAL---PEMEQEQRTELSADKEFRATCTDSTISIL 1203
QY 368 TEYKRRKQOHNLSAFPKEVKEKKGEPLPGQDPSPLQRPQHLMDQGMK----- 418
DB 1204 DQESWRD---SASP-----DAGEGLGRPESS---QKAIREAQMGQNERPWASSLT 1250
QY 419 HSPSAGPELLAQDKRPRSGTSSSVSVR 448
DB 1251 HSPESHPLCTLQHQERDESTLATSLEKQVR 1280

RESULT 8
US-10-106-698-5295
; Sequence 5295, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: P4005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 5295
; LENGTH: 557
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-106-698-5295

Query Match 3.7%; Score 100; DB 15; Length 557;
Best Local Similarity 21.0%; Pred. No. 1.9;
Matches 89; Conservative 66; Mismatches 171; Indels 98; Gaps 17;

QY 120 RTAREPDTSPGLCKLKKVSGIGAAANLYR-----QASMFNIYFHALVCAVLTN 169
DB 7 RSRRMGPNQRLNGPOGRRTKEGGAASGLRLHTRAPGEGAM---LWFQAI PAIAT 63
QY 170 QETITAEQYKVLFEDEDERSTDSSQCCSDEDEIFETTAQVSPPRKEKQWRAMPPLS 229
DB 64 AKSGAVFVFA--GDEQST---QMAASWEDKVTASNSFVALKIDTKSEACLOFSQ 119
QY 230 VQPV-----SNADWWLVYKRLHKLCEMLCNNYIQML----- 261
DB 120 IYVVCVPSFPFIQDGGIPLVLAAGVSAD--ELVYRIHV-----RQMLHKSETS 169
QY 262 -----DLNCEMEBPPIFKGDPFILPFSQSESTPSTGSGKETPS 303
DB 170 VANGQSESSVSTPSPASFEBNNTCENQSRNAELCEIPSTDYKSDTATGSGSAGHATSS 229
QY 304 EDRSQRHEMGSLSLKAGGDLPLPSPKVEKKDPSRKKEMWENAGNKITYMAADKTI 363
DB 230 QEPSGCSDDQPAEDLNIRV---ERL---TKLEBRREKRR---EEQREIKTEIRKKT 280
QY 364 SKLMTYKRRKQOHNLSAFPKEVKEKKGEPLGPRGQDPSPLQRPQHLMDQGMHFSFA 423
DB 281 GKEMLDY--KXQEBELTKRMLERNRKADPRARER-----IKQIALBRAEPAARFA- 333
QY 424 GPBLLRQDKRPRSGTSSLSVSRDAEQIQAMTNVTLVNLQIQI-LPD-QTFALQ 481
DB 334 -----KTKEVEAAKAAALLAKQAEWEVRESYARERSIVA-RIQRLPDGSSFTNQFP 386
QY 482 AVFP 485
DB 387 SDAP 390

RESULT 9
US-09-788-600-5
; Sequence 5, Application US/09788600
; Patent No. US20020004489A1
; GENERAL INFORMATION:
; APPLICANT: Shi et al.
; TITLE OF INVENTION: Retinoid Receptor Interacting Polynucleotides, Polypeptides, and
; FILE REFERENCE: PTO17P1
; CURRENT APPLICATION NUMBER: US/09/788,600
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: PCT/US00/22351
; PRIOR FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: 60/189,026
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/148,757
; PRIOR FILING DATE: 1999-08-16
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-788-600-5

Query Match 3.6%; Score 98; DB 9; Length 411;
Best Local Similarity 19.1%; Pred. No. 1.8;
Matches 74; Conservative 54; Mismatches 115; Indels 144; Gaps 17;

QY 2 IRAMAQOVFMILDTQCSPTKPTNPFDAQSCOLIIELEPDEKPNHTKKSVSFREIVVSLLS 61
DB 73 IKSUKETLLER---EFTT-----SHQSSQGIYVETSEE-----GNSVASQSVAAITLS 119
QY 62 HQVLLQNLVDILLEEFVKGPSPEEKTIQVPEAKLAGFLRYISMQNLAVIFDLLDVRYT 121

Db 120 KRSLL-----VLMPE-----SSABETIV-CPETOLSS-----SETPDL----- 150
Qy 122 AAEFTSPGLKCLLKV-----SGIGAAALYROSAMS-----FN 156
Db 151 --EREVSPPSRDILDOVRILIMADKEVGNKEDAEKVAISTFSSNQVSCPLCQCQFPPTK 208
Qy 157 IYFHALVCAVLNTOEITTEQYKVLFFEDERSTOSSQCCSEDEDIPEETNOVSPPRGK 216
Db 209 IERHAMVCGMLBEDVLTRRQKEAKTKSDSGT-----AAQTSIDIDK 251
Qy 217 EKORARAPMLISVQPVSNADWWVLVKRLKLMELCNNTI-----OMHLDLNCMBEPP 271
Db 252 NEKKY-----LCKSLVFPREYQCHVD--SLQLAK 279
Qy 272 IFKGDPFLLPFSQESSSTPTGSGSK-----ETPSEDDRSQSR-----EHMGSL 318
Db 280 ADQGD-----GEGSGRACSTVEGKQGRLLKPKKEKHSEGRLLSFLQSEHKTSDA 331
Qy 319 SLKAGGDDLPLPSPVKEKKDPSRKKE 345
Db 332 DIKSETGAFRVSPGMEERAGCSREWQ 358

RESULT 10
US-09-738-626-4700
; Sequence 4700, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASARU
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4700
; LENGTH: 446
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4700

Query Match 3.6%; Score 98; DB 10; Length 446;
Best Local Similarity 21.1%; Pred. No. 2;
Matches 96; Conservative 59; Mismatches 149; Indels 152; Gaps 23;
Db 125 FDTSPGLKCLLKVSGIGAAALYROSAMSFNIFYHALVCAVL-----NTOEITTA 176
Qy 31 FEIAGEERILLTGAGAGAGSTLL-----ALAGVLGSGDSGVSTGELLVDP 77
Db 177 QYKVLFFEDDER-----STDSQOCS--EDEDIFEETAO-----VSPPR-- 214
Qy 78 SIGVLVQDPDSQVIAIRIGDDVAFGCGENQIPREIIMPRVERALVELVGLDPLSHPTKYL 137
Db 215 --GKEKR-----QWRAMPILSVQPVSNAD-----WWLVKRL 245
Qy 138 SGGOKORLALAGVIAWGATLILID--EPTANLDPQGGKDVAAVDRVOETGATLIVVEHR 196
Db 246 HLLCMELCNNTYIQMHLDLNCMBEPPPIFKGDPFLLPFSQESSSTP-----STGGF 296

Db 197 HELWNIIDRIISL-TGEDVQPAELIKVQ-----LGAOPSTISKPLIMANDLCTWGL 251
Qy 297 SGEKTPSEDDRSQSRHNG-----ESLSLAKAGGDDLPLPSPVKEKKD-----PSR 342
Db 252 RSFEVP-----EGASTVITGPRGAKSLTALTMGS--LLPPRSGLLELSDTVRGGLNTPPH 305
Qy 343 KKEWMENA--GNKIYMAADKTISKLMTEYKQRKQHNLSAFPPEVYVEKKGEPLGPRGQ 400
Db 306 K--WRSADLAARIGTFVQDPE-----HQFVARTVRDELEIGPKIMKVDAS----- 348
Qy 401 DSPILQROHMDGOMHSPSAGPELLRODKRRSGSTSSLSVSRDAEAQIANTNM 460
Db 349 -----ERLELDLDRLLHLENANPFTL-----SGEKRRLSVAT----- 383
Qy 461 VLTVNLQIQLPDQFTALQPAVPCISOLTCVTD 496
Db 384 ALVAAPKLLIDEPFL--GQDEFTTELVTMRELTD 418

RESULT 11
US-09-982-091A-4
; Sequence 4, Application US/09982091A
; Patent No. US20020151030A1
; GENERAL INFORMATION:
; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY
; APPLICANT: KUMAGAI, AKIO
; APPLICANT: DUNPHY, William
; TITLE OF INVENTION: CLASPIN PROTEINS AND METHODS OF USE THEREOF
; FILE REFERENCE: CIT1320-1
; CURRENT APPLICATION NUMBER: US/09/982,091A
; CURRENT FILING DATE: 2002-10-17
; PRIOR APPLICATION NUMBER: US 60/241,246
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-982-091A-4

Query Match 3.6%; Score 98; DB 10; Length 1332;
Best Local Similarity 20.4%; Pred. No. 11;
Matches 73; Conservative 53; Mismatches 124; Indels 108; Gaps 15;
Qy 168 TNOEITTAQYK--KVLFFEDERSTDSQCCSEDEDIPEETAQVSPPRGKQRAR 225
Db 46 SDEIFVSKKLKNNKVLQDSSETEDTNASEKTTYSABEENKENLYAGNNTK----- 99
Qy 226 PLISVQPVSNADWWVLVKRLKLMELCNNTYIQMHLDL-----CMBEPPPIFKGDPF 278
Db 100 -----IKRIYTVADSDSEYMSLQENLEAQVKPCL----- 133
Qy 279 FILPFSQESSSTPTGSGKETPSEDDRSQSRHNG-----GESLSLAKAGGDDLPLPSPK 334
Db 134 ---LSLQSGNSTDT-----TDKSSKKHIDKGTAKAKAYKSRLEKEERK 179
Qy 335 VEKQDPSRKKEW-----WENAGNKIYMAADKTISKLMTEYKQR--KQOHNL-----SA 381
Db 180 MEKIRQLKKETKQGEDVDEQPFNDSCGLVDKDLFTGLJEDBNNSPLEDBESLESIRAA 239
Qy 382 FPKVYKVEKGEPLGPRGQDSPILQROHMDGOMHSPSAGPELLRODKRRPSGSGTS 441
Db 240 VKNVYKHKKKKP-----SLESQV--HSPEEGSELKGTTRKERRKARL 281
Qy 442 SLASVDAEAQIANTNMVL--TVNLQIQLPD-----QFTALQPAVPCISOLTC 493
Db 282 S-----KEALKQHSSTORLIRBSALNDPYHMPENKTIHDPFKKP-----RPTCH 327

RESULT 12
US-09-823-187-88


```

; Sequence 88, Application US/09823187
; Publication No. US20030096952A1
; GENERAL INFORMATION:
; APPLICANT: Burgees, Catherine
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Liu, Xiaohong
; APPLICANT: Majumder, Kumud
; APPLICANT: Padigaru, Muradilhar
; APPLICANT: Patutajan, Meera
; APPLICANT: Shimkets, Richard A
; APPLICANT: Spaderna, Steven K
; APPLICANT: Spytek, Kimberly
; APPLICANT: Taudier, Raymond J
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-745
; CURRENT APPLICATION NUMBER: US/09/823,187
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/193,339
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/193,205
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/195,343
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: 60/195,088
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,005
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,792
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: 60/196,556
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: 60/197,081
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/197,525
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/197,087
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 103
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 88
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-823-187-88

Query Match          3.6%; Score 97.5; DB 11; Length 669;
Best Local Similarity 18.6%; Pred. No. 4.3;
Matches 79; Conservative 57; Mismatches 125; Indels 163; Gaps 18;

```

```

QY 473 DQFTALQ---PAVPCISQL-----TCHWTD-----IRVQAVREMLGRV 510
;
;
;
Db 476 SEIKFALKVNDPVRKCLSLAEELGLQVTSLLQKNVDVATLKIRRYKAMDVAKA 535
QY 511 GRVY 514
;
Db 536 AEVY 539

RESULT 13
US-09-764-864-823
; Sequence 823, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT23
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PAM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 823
; LENGTH: 759
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (18)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (19)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (21)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (257)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (239)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-764-864-823

Query Match          3.6%; Score 96.5; DB 10; Length 759;
Best Local Similarity 19.1%; Pred. No. 6.4;
Matches 101; Conservative 84; Mismatches 184; Indels 159; Gaps 26;

```

```

Db      612 CHLYSFCLOKCEVTEEGQTRWTCYKCSSNKV-----GKLSENSEIKGR----- 660
Qy      380 SAPEKVKV-----EKKEPLGPRGQDSPLORPOLMDQGMHRSFSAGPEL-----LR 429
Db      661 -ITFQVKKSPSYHOSKQDPTAKKTSEPLVD-PQIQAFDQLCLYRSSSLALLTELS 718
Qy      430 QDK-----RPRSGTSGSSLSVVRDAEAOIQMTMVLTVLNOIQILP 472
Db      719 QNRSSSEYRPFSGSOSAPAFNSIFQNE-----NF-----QLOQLIP 753

```

RESULT 14

```

US-09-764-864-1161
; Sequence 1161, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; PRIOR FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 1161
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-864-1161

```

Query Match 3.6%; Score 96; DB 10; Length 555;

Best Local Similarity 21.0%; Pred. No. 4.4; Indels 82; Gaps 14;

Matches 64; Conservative 46; Mismatches 113; Indels 82; Gaps 14;

```

Qy      192 SSQOCS-----SEDEDIFEETAQVSPPRGKE---KROWARMPILSVQVSNADWVW 240
Db      3 SMLQCSMKKKVLSDEDEKADVPCTSTRKKDHQRRRLRNAAQSYDIO-----AW 55
Qy      241 LVKRLHKLCEMLCNVYIQMHLLENCEBEPPIFKGDPFILPSFQSSSTPSTGSGSKE 300
Db      56 -----KKQCEBLLNLIFQ-----CEDSEFPQPVLDLEYDYRDIIDTPM----- 95
Qy      301 TPSEDDRQSRHMGESLSLKAAGDILLPSPKVEKDP-----SRKKEMWENAGNKIYT 356
Db      96 -----DFATVRE-----TLEAGNYE-----SPMELCKDVLLIFNSKAYTPSKRSRIYS 139
Qy      357 MAA-----DKTISKLTMEYKKRQOHNLSAPPEKVEKKEGPELGRGQDSPLORPOH 410
Db      140 MSLRLSAFEEHISVLSYKALRFHKRNTTKRRKKRNRSSSVSSAASP--ERKKR 197
Qy      411 LMDQGMHRSFSAGPELLQDKRPRSGSTGSSLSVVRDAEAOIQAMT---NMVLTVLNO 467
Db      198 ILK-----PQL--KSESSSTAFASTPTRSIPPRHNAQINGKTESSSVVRTRSNR 244
Qy      468 IQILP 472
Db      245 VVVDP 249

```

RESULT 15

```

US-09-764-864-1574
; Sequence 1574, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; PRIOR FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0

```

```

; SEQ ID NO: 1574
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-864-1574

```

Query Match 3.6%; Score 96; DB 10; Length 555;

Best Local Similarity 21.0%; Pred. No. 4.4; Indels 82; Gaps 14;

Matches 64; Conservative 46; Mismatches 113; Indels 82; Gaps 14;

```

Qy      192 SSQOCS-----SEDEDIFEETAQVSPPRGKE---KROWARMPILSVQVSNADWVW 240
Db      3 SMLQCSMKKKVLSDEDEKADVPCTSTRKKDHQRRRLRNAAQSYDIO-----AW 55
Qy      241 LVKRLHKLCEMLCNVYIQMHLLENCEBEPPIFKGDPFILPSFQSSSTPSTGSGSKE 300
Db      56 -----KKQCEBLLNLIFQ-----CEDSEFPQPVLDLEYDYRDIIDTPM----- 95
Qy      301 TPSEDDRQSRHMGESLSLKAAGDILLPSPKVEKDP-----SRKKEMWENAGNKIYT 356
Db      96 -----DFATVRE-----TLEAGNYE-----SPMELCKDVLLIFNSKAYTPSKRSRIYS 139
Qy      357 MAA-----DKTISKLTMEYKKRQOHNLSAPPEKVEKKEGPELGRGQDSPLORPOH 410
Db      140 MSLRLSAFEEHISVLSYKALRFHKRNTTKRRKKRNRSSSVSSAASP--ERKKR 197
Qy      411 LMDQGMHRSFSAGPELLQDKRPRSGSTGSSLSVVRDAEAOIQAMT---NMVLTVLNO 467
Db      198 ILK-----PQL--KSESSSTAFASTPTRSIPPRHNAQINGKTESSSVVRTRSNR 244
Qy      468 IQILP 472
Db      245 VVVDP 249

```

Search completed: July 24, 2003, 12:18:14

Job time : 63.1151 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: July 24, 2003, 11:48:38 ; Search time 42.4715 Seconds
(without alignment)
1172.914 Million cell updates/sec

Title: US-09-991-681-27
Perfect score: 2698
Sequence: 1 RIRAAQVFMIDTQCSPKT.....VRQAVREWLGRGVYDIIV 518

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	126	4.7	2278	1 S56274	FAB1 protein - Yea
2	124	4.6	2150	1 S27802	zinc finger protei
3	124	4.6	2150	2 T19450	hypothetical prote
4	113	4.2	482	2 T22981	hypothetical prote
5	111	4.1	781	2 T41551	hypothetical prote
6	111	4.1	1148	2 T09073	splicing factor Sl
7	110	4.1	829	2 A33166	colorectal tumor s
8	109.5	4.1	587	2 T24103	hypothetical prote
9	108.5	4.0	2783	1 A41948	alpha-fetoprotein
10	108	4.0	506	2 T41788	global transactiva
11	108	4.0	1173	2 T25985	hypothetical prote
12	108	4.0	1912	2 T29088	vitellinogenin I pre
13	106.5	3.9	2617	2 AE2136	peptide synthetase
14	105	3.9	506	2 B72855	global transactiva
15	104.5	3.9	830	2 T16236	hypothetical prote
16	103.5	3.8	1304	2 T19397	hypothetical prote
17	103	3.8	453	2 T22153	hypothetical prote
18	103	3.8	1306	2 T13592	hypothetical prote
19	103	3.8	1445	2 T15212	hypothetical prote
20	103	3.8	1502	2 T42216	multidrug resistan
21	102.5	3.8	439	2 B84774	hypothetical prote
22	102	3.8	457	1 A28468	chromogranin A pre
23	101.5	3.8	621	1 S59632	endo-1,4-beta-xyla
24	101.5	3.8	682	2 B84021	two-component sens
25	101.5	3.8	2101	2 A42184	nuclear mitotic ap
26	101	3.7	859	2 T01461	hypothetical prote
27	101	3.7	1099	2 A59311	myosin VII1, ZMM3
28	101	3.7	1198	2 T20262	hypothetical prote
29	101	3.7	1203	2 C89217	protein C55A6.2 [1

30	100.5	3.7	453	2 S50248	SSF2 protein - Yea
31	100.5	3.7	484	2 S60943	RUD3 protein - Yea
32	100.5	3.7	947	2 G70657	probable adi - Myc
33	100	3.7	1268	2 A44125	high density lipop
34	99.5	3.7	342	2 A56552	homeotic protein H
35	99.5	3.7	649	2 AG0710	DNA topoisomerase
36	99.5	3.7	1534	2 A56734	ribosome receptor,
37	99	3.7	456	2 S62962	hypothetical prote
38	99	3.7	822	2 T41622	probable ABC trans
39	99	3.7	1207	2 D84781	hypothetical prote
40	98.5	3.7	450	2 T21931	hypothetical prote
41	98.5	3.7	739	2 T21431	hypothetical prote
42	98.5	3.7	960	2 S44812	F44B9.6 protein -
43	98.5	3.7	986	2 T41809	DNA polymerase orf
44	98.5	3.7	1750	2 E86151	hypothetical prote
45	98	3.6	692	2 E96841	hypothetical prote

ALIGNMENTS

RESULT 1

S56274 FAB1 protein - yeast (Saccharomyces cerevisiae)

N/Alternate names: protein YFR019w

C/Species: Saccharomyces cerevisiae

C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000

C/Accession: S56274; A55997

R/Murakami, Y.; Naitou, M.; Hagiwara, H.; Shibata, T.; Ozawa, M.; Sasamura, S.I.; Sasaau

submitted to the EMBL Data Library, May 1995

A/Description: Analysis of the nucleotide sequence of chromosome VI from Saccharomyces ce

A/Reference number: S56186

A/Accession: S56274

A/Molecule type: DNA

A/Residues: 1-2278 <MIR>

A/Cross-references: EMBL:D50617; NID:9836685; PIDD:BA009258.1; PTD:9836774; GSPDB:GN0000

R/Yamamoto, A.; Koshland, D.

submitted to GenBank, August 1993

A/Description: FAB1: a novel gene required for endocytic-vacuolar pathway and nuclear mi

A/Reference number: A55997

A/Accession: A55997

A/Molecule type: DNA

A/Residues: 1-2274, 'R', 2276-2278 <YAM>

A/Cross-references: GB:U01017; NID:g392993; PIDD:AAA81360.1; PID:g398498

C/Genetics:

A/Gene: SCD:FAB1; MIPS:YFR019w

A/Cross-references: SCD:S0001915; MIPS:YFR019w

A/Map position: 6R

C/Superfamily: FAB1 protein, 1-phosphatidylinositol-4-phosphate 5-kinase homology <PIK>

F/1911-2266/Domain: 1-phosphatidylinositol-4-phosphate 5-kinase homology <PIK>

Query Match 4.7%; Score 126; DB 1; Length 2278;

Best Local Similarity 21.5%; Pred. No. 0.98; Matches 108; Conservative 77; Mismatches 148; Indels 170; Gaps 27;

QY	19	KTPNNF--DHAQSC--QLIIEL-----PPDE--KENGTKKVSFREIVS	58
DB	1330	KCPHDFKDKVYKGVNDIVRLEYSDELVEHLLTPRKIKMKRHIDIKLVE-----	1382
QY	59	LSHQVLT--LQNTYDILLE--EPYKSPSPGEKTIQVPEAKLACFLYISMONTAVIFDL	114
DB	1383	LVYKILIEKINNFSGVLSRLERIKLDSMTKDKV-----LSQAKIIEIKSNATEEQK	1434
QY	115	L-----LDSYRTARFDPSPGKCLKLKVSIGSGANLYROSAMSFNYFAHLVCAV---	167
DB	1435	LMQDLDTFYADSPCDQHLPLNLVTK-----SLY-DKAVVWNSSTFAIPAKSYLPSE	1484
QY	168	TNCFITAEQVKVLFEDDERSTDSSQCSSEDEDIPEETAQVSPPGKEKRGQRAMPPL	227
DB	1485	TDISRTIRAKQKGLFY-----DSSRKDSBDKSLNDEKAKTRKPKENE-----LPL	1530
QY	228	LSVQPSNADVMVLVLRHLKLCMELCNVYIQMHLDLENCE-----EP-----PIFKG	275

Db 1531 EGLKQVE-----KPKIDSKTTEKRDPTNEQNAVITTTKCD 1567
 Qy 276 DPFLLPFPQSE--SSTPTGSGFKETPSEDDRSQ--SREHMESLSLKAGGDLPLPSP 333
 Db 1568 DTPILPTGSHLTPTPSASVSSSLTPQTEPRPPIRSRGTCISMT----- 1613
 Qy 334 KVEKDPKPKKEMWENAGKIYTMADKT---ISLTMTEYKKRKQ-----QHNLISA 381
 Db 1614 ---HDKSTRP-----NIRKMSSDSLGLAGLALNEYSKNNVSKLAFEDDMHFDPA 1661
 Qy 382 FPKKVEKKEKKEPL-----GPRGDS-PLQRPQHLMDQGMRSFSAPELLRODK 432
 Db 1662 LSKKELERERERLQNKDKYIRLQSTPIVEIYKAVKD-----AYDE 1706
 Qy 433 RPRSGTSSSL-SVSVRAEAQI 454
 Db 1707 PLHSRSSGNLSSANVKLEAPV 1729

RESULT 2

S27802
 zinc finger protein - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 08-Dec-2000
 C:Accession: A45172; S27802
 R:Klein, R.D.; Meyer, B.U.
 Cell 72, 349-364, 1993
 A:Title: Independent domains of the Sdc-3 protein control sex determination and dosage
 A:Reference number: A45172; MUID:93161411; PMID:8431944
 A:Accession: A45172
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-2150 <KLE>
 A:Cross-references: EMBL:M85149; NID:G156440; PIDN:AAA28144.1; PID:G156441
 A:Experimental source: strain N2
 A:Note: sequence extracted from NCBI backbone (NCBI:P124842)
 C:Genetics:
 A:Gene: sdc-3
 A:Intron: 70/1; 102/1; 153/3; 258/1; 305/1; 1283/1; 1398/1; 1703/2; 1853/1; 1887/1; 190
 C:Superfamily: zinc finger protein sdc-3
 C:Keywords: DNA binding; nucleus; zinc finger

Query Match 4.6%; Score 124; DB 1; Length 2150;
 Best Local Similarity 20.6%; Pred. No. 1.3; Mismatches 166; Indels 210; Gaps 24;
 Matches 114; Conservative 63; Mismatches 166; Indels 210; Gaps 24;
 Qy 33 IIELEPPDEKPNHTRKKSVPREIVVSL-----SHOVLQ-----NLVDILLEEFVK 79
 Db 1177 IERLSEDDQLDQGSSEDISFEIIVDLLESGVEVQVNVQSVTISRSTSPFSLVEDPBE 1236
 Qy 80 GP-----SPGEKTIQ-VPEAKLAGFLRYISQNLAVIFDILLDSYRTAREPDTSPGLKC 133
 Db 1237 HPEQLPVASASEKANNQIVPEVEVGSVPVTVNQ-----EENVTSSEG--P 1279
 Qy 134 LKKVSGIGAGANLYRQSGAMSFNIYFHALVCVLTNQ-----ETITAEQ----- 177
 Db 1280 TLQEGSSIPSSSHY-----TVDELLGTESPGREATETTPVAESPPKKSGKT 1326
 Qy 178 -----VKVLFEDDERSTDSQQCSSEDEDIPEETAQVSPRGEKQKQRMPL--L 228
 Db 1327 TRGRPKVKYKKNLKKRIQRRGQKEBAHEPE-VVEEQGVPEVPEVHVPVAPAAQL 1385
 Qy 229 SVQPVSNADWMLVKRLHKLKMLCNNTYIQHMLDLENMEEP-----PIFGDPPFLLPS 283
 Db 1386 ETEPT-----EQOIEEPDKVFEPPIEA-----LPL 1410
 Qy 284 FOSSESTPTSGFGSKETPSEDD-----RSQREHMG 315
 Db 1411 FETSPVAPPEGNITPRAHSSDDVQVVISSETDPNGPILNVEQVQNDKLTAYQYSTEEILG 1470
 Qy 316 ESLSLKAGG-----DLLPSPKVEKQDPSRKKEWMEN--AGNKIYTMAA 359
 Db 1471 EYGLDEAGAPSPSEIIVHDEVLDQVLPNPKSKKGRGRKKTTPPHIAKARKVFT--- 1527

Qy 360 DKTISKLMT-----YKKRQOHNLSA-----FPKEVYKKEGEP-----LQPR 398
 Db 1528 --SISK--TEIEELAPPTQOSRKRMANVSSEEAATATRRQKRAKEEENDDVSRVLTPE 1583
 Qy 399 GQSPFLQRPQHLMDQGMRSFSAPELLAQDPRRSGSGSSLSVS----- 446
 Db 1584 PEDLHETRPQGVGEGETPSTLRTGRE-----STASSVKTISRSKRLFLSKNP 1632
 Qy 447 VRDAEAQIQAMTN 459
 Db 1633 VPRMRIOQAGTN 1645

RESULT 3

T19450
 hypothetical protein C25D7.3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 11-Jan-2000
 C:Accession: T19450
 R:Ainscough, R.
 submitted to the EMBL Data Library, October 1996
 A:Reference number: Z19125
 A:Accession: T19450
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-2150 <WLL>
 A:Cross-references: EMBL:281039; PIDN:CA802774.1; GSPDB:GN00023; CESP:C25D7.3
 A:Experimental source: clone C25D7
 C:Genetics:
 A:Gene: CESP:C25D7.3
 A:Map position: 5
 A:Intron: 70/1; 102/1; 153/3; 258/1; 305/1; 1283/1; 1398/1; 1703/2; 1853/1; 1887/1; 190
 C:Superfamily: zinc finger protein sdc-3

Query Match 4.6%; Score 124; DB 2; Length 2150;
 Best Local Similarity 20.6%; Pred. No. 1.3; Mismatches 166; Indels 210; Gaps 24;
 Matches 114; Conservative 63; Mismatches 166; Indels 210; Gaps 24;
 Qy 33 IIELEPPDEKPNHTRKKSVPREIVVSL-----SHOVLQ-----NLVDILLEEFVK 79
 Db 1177 IERLSEDDQLDQGSSEDISFEIIVDLLESGVEVQVNVQSVTISRSTSPFSLVEDPBE 1236
 Qy 80 GP-----SPGEKTIQ-VPEAKLAGFLRYISQNLAVIFDILLDSYRTAREPDTSPGLKC 133
 Db 1237 HPEQLPVASASEKANNQIVPEVEVGSVPVTVNQ-----EENVTSSEG--P 1279
 Qy 134 LKKVSGIGAGANLYRQSGAMSFNIYFHALVCVLTNQ-----ETITAEQ----- 177
 Db 1280 TLQEGSSIPSSSHY-----TVDELLGTESPGREATETTPVAESPPKKSGKT 1326
 Qy 178 -----VKVLFEDDERSTDSQQCSSEDEDIPEETAQVSPRGEKQKQRMPL--L 228
 Db 1327 TRGRPKVKYKKNLKKRIQRRGQKEBAHEPE-VVEEQGVPEVPEVHVPVAPAAQL 1385
 Qy 229 SVQPVSNADWMLVKRLHKLKMLCNNTYIQHMLDLENMEEP-----PIFGDPPFLLPS 283
 Db 1386 ETEPT-----EQOIEEPDKVFEPPIEA-----LPL 1410
 Qy 284 FOSSESTPTSGFGSKETPSEDD-----RSQREHMG 315
 Db 1411 FETSPVAPPEGNITPRAHSSDDVQVVISSETDPNGPILNVEQVQNDKLTAYQYSTEEILG 1470
 Qy 316 ESLSLKAGG-----DLLPSPKVEKQDPSRKKEWMEN--AGNKIYTMAA 359
 Db 1471 EYGLDEAGAPSPSEIIVHDEVLDQVLPNPKSKKGRGRKKTTPPHIAKARKVFT--- 1527
 Qy 360 DKTISKLMT-----YKKRQOHNLSA-----FPKEVYKKEGEP-----LQPR 398
 Db 1528 --SISK--TEIEELAPPTQOSRKRMANVSSEEAATATRRQKRAKEEENDDVSRVLTPE 1583
 Qy 399 GQSPFLQRPQHLMDQGMRSFSAPELLAQDPRRSGSGSSLSVS----- 446

Db 1584 PEDLETERPGHVEEGFETPLRTGRB-----STASSVKTSSRSKRLFLSKNNP 1632
 Oy 447 VRDAEOIQAWTN 459
 Db 1633 VPRMRISQACTN 1645

RESULT 4

T22981
 hypochlorite protein F59B10.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T22981
 R:Lloyd, C.

submitted to the EMBL Data Library, March 1995

A:Reference: T22981
 A:Accession: T22981

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-482 <WIL>

A:Cross-references: EMBL:Z48716; PIDN:CAA88601.1; GSPDB:GN00020; CESP:F59B10.2

A:Experimental source: clone F59B10

C:Genetic8:

A:Gene: CESP:F59B10.2

A:Map position: 2

A:Introns: 40/2; 74/1; 118/3; 415/1

Query Match 4.2%; Score 113; DB 2; Length 482;

Best Local Similarity 21.9%; Pred. No. 1;

Matches 85; Conservative 46; Mismatches 128; Indels 130; Gaps 16;

Oy 118 SYTAREFDTS-----PGLKCLTKKVSIGGAANLYROSAMSFNIFHALVCAVLTN 169
 Db 81 SLKQIRELRNSIDSLIKPELKLKLRKAA-----EESKRAHFESIOEEKIS 130
 Oy 170 QETITAEQVKVLFEDERSTDSQCSSEDEDFEETAQ---VSPRGKKEKQWRAMP 226
 Db 131 EDLPTAQLPSKSPF---KKAIGSSSSSDSIIFDEVEFLVPSPP----- 175
 Oy 227 LLSQVPSNADWVLVRLKLCHELCNNYIQMLDLNCEHEPPI-----FKGDPF 278
 Db 176 -KPAAPARTAVIV-----VEKLEKPAVVEQKARKKKEKPT 210
 Oy 279 FILSPFQ-----SESTPSTGFGSGKETPESEDR---SOSREHGESLSLKAGGD 326
 Db 211 PTSEFSSSSSSSTSESTSESSSSSSSESSSSSSSQVSSKTSKSSSKAYGSD 270
 Oy 327 LLLPPSPKVEKDDPSRKKEWMENAGNKIYTMAADTKISKL---MTEYKRRKQOHLSAF 382
 Db 271 F-----ESEKSSSS-----SASTISKVTPKKLDKPKQTKKPKDKRAK 307
 Oy 383 PKVEKVEKGEPLPRGDSPLQRPQHLMQCGMRHSFSGPEL----- 428
 Db 308 PDDIRONKKEPPI-PYEDFIPKLSSRSSNSSEKSTVRETNRTLESKLTKINKSSEKME 366
 Oy 429 --RODKR--PRSGTSSLSVSRDAEAO 453
 Db 367 KPRKDIRAPRSSSSSS---TLRDAERE 392

RESULT 5

T41551
 hypochlorite protein SPC70.05c - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C:Accession: T41551
 R:Wood, V.; Rajandream, M.A.; Bartell, B.G.; Seeger, K.; Harris, D.

submitted to the EMBL Data Library, June 1998

A:Reference number: Z22001

A:Accession: T41551

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-781 <WOO>

A:Cross-references: EMBL:AL023794; PIDN:CAA19355.1; GSPDB:GN00068; SPDB:SPC70.05c
 A:Experimental source: strain 972h-; cosmid c70
 C:Genetics:
 A:Gene: SPDB:SPC70.05c
 A:Map position: 3

Query Match 4.1%; Score 111; DB 2; Length 781;
 Best Local Similarity 21.4%; Pred. No. 2.9;
 Matches 75; Conservative 41; Mismatches 116; Indels 118; Gaps 19;

Oy 189 STSSQCCSEDEDFEETQVSPPRKKEKQWRAMPPLSVQPVSNADWVLVRLKHL 248
 Db 74 TTDSMAHPVGNNSPL-QTPQKSPPRKH-----TAPAPPI-PVS-----ASRHHK- 116
 Oy 249 CMELCNNYIQMLDLNCEHEPPIFKGDPFPLPSFSESTSTST-----GSGF----- 296
 Db 117 -----PHSGKLKLEKA-----MPPGKSNANSPTSESPSKGFSFINNH 158
 Oy 297 -----SGKETPESEDRSOSREHGESLSLKAGGDLLPPSPKVEKQPSRK 343
 Db 159 LHKQTSHPSPVNGKS--SDHKSQSYOHLKNS-----PNSRTARKPVPRR 204
 Oy 344 KEMMENAGNKIYTMAADTKISKLMTEYKRRKQOHLSAFPEYK--VEKGEPLGPRGQ- 400
 Db 205 -----ANSASHNLGSTRKSPGNAKESLSRSABLPKAKEMETNGYRKPSPLSPNSJ 258
 Oy 401 -----DSPLQRPQHLMQCGMRHSFSGAPELQDPRPRSGSTSS----- 442
 Db 259 RNEGNGSYFDGDLTSPSPSSPTG-TNPSMSKSPSL-----SSLASTGASTRPGPSK 311
 Oy 443 -LSVSRD--AEAQIQAWTN-----MVLTVLNOIQILPDOT-FTALQPA 482
 Db 312 PIVSRVADNVANNTSYESMHPSTFEDMTVAVSSSLKLTPTGTGDCINPA 361

RESULT 6

T09073
 Splicing factor Sipl - human

C:Species: Homo sapiens (man)

C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000

C:Accession: T09073

R:Zhang, W.J.; Wu, J.Y.
 Mol. Cell. Biol. 18, 676-684, 1998

A:Title: Sipl, a novel RS domain-containing protein essential for pre-mRNA splicing.

A:Reference number: Z16554; MUID:98107652; PMID:9447963

A:Accession: T09073

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1148 <ZHA>

A:Cross-references: EMBL:AF030234; NID:G2822459; PIDN:AAC39565.1; PID:G2822460

A:Experimental source: HeLa cells

C:Keywords: pre-mRNA splicing

Query Match 4.1%; Score 111; DB 2; Length 1148;
 Best Local Similarity 18.5%; Pred. No. 5.1;
 Matches 94; Conservative 76; Mismatches 183; Indels 154; Gaps 20;
 Oy 1 RIRAMAQVMDLTQSPKTPNNFPAHOSCOLIIELP---PDEKPNHTKSSVFRITIV 57
 Db 213 QISGLSQSEVKTIV-CTVHLPNDFPTCLTSESRYQVSCPLDSLSENVSVAEEKITE 271
 Oy 58 SLLSHOVLQNLVDILIEEFVKGP---SPGE-----EKTIOVPEAKLAGFLRYISMON 107
 Db 272 SSLV-EITTEKDPFLTKTEELIESPKLESSGCEIITQYDROSIVSPVQLGHVETDEVEI 330
 Oy 108 LAVIFDILLDSYTRAREFDTSPGLKCLKKVSIGGAANLYROSAMSFNIFHALVCAVL 167
 Db 331 IATCDTRGNDFNNIIQGSENN-----LLKN-----NLNLTGLKSLK 367
 Oy 168 TNOETIT-----AEQYKVLFEEDER---STDSSQCCSSEDEDFEETQVSP 213
 Db 368 EKNESTLHPRSTELPKTHIEIQIKHSEEDNNEMIPWECDSFCSQDNE----- 416

```

QY 214 RGKEKQWRAMPPLSLVQPVSNADWVLVKRLHKLQMLCNNTYQMLHLDLENCEPPIIF 273
DB 417 -----EVEPSVAD-----LKONNENSVTHCS-----ENNM--PSSD 446
QY 274 KGDPEFIIIPFQSES-----STPSTGFGSGKETPEDDRSQSREHMG 315
DB 447 LADEKVEIVSQSSESPKOTIDTKKPRRRSRFHSPTTWSPNKOTPOEKRRPOS----- 501
QY 316 ESLSLKAGGDLILPPSPKVEKKDPSRK-----KEMWENAGNKITYMAADTKISKLMTE 369
DB 502 -----PSPRERETGKESRKSSQSPSPKESARGRKRSQSPKKDIA----- 541
QY 370 YKKRQOHNLSAFPKEVVEKKEGLSPRGQSPILQRPQHMDGQMRHSFAGPELLR 429
DB 542 -RERKQOSRSPKRTTRESRSSELSPPRETS-----RENKRQOPRVKDS-SPEKRSRS 594
QY 430 QDKRPRSGSTGSSLSVSVADAQIOAWT 458
DB 595 QSRERESDRDQGR-----RERERRRRKMS 618

```

RESULT 7

```

A33166
colorectal tumor suppressor protein - human
C.Species: Homo sapiens (man)
C.Date: 12-Apr-1991 #sequence_revision 12-Apr-1991 #text_change 21-Jul-2000
C.Accession: A38434; A33166
R.Kinzler, K.W.; Nibert, M.C.; Vogelstein, B.; Bryan, T.M.; Levy, D.B.; Smith, K.J.;
Science 251, 1366-1369, 1991
A.Title: Identification of a gene located at chromosome 5q21 that is mutated in colorec
A.Reference number: A38434; MUID:91164855; PMID:1848370
A.Accession: A38434
A.Status: preliminary
A.Molecule type: mRNA
A.Residues: 1-829 <KIN>
A.Cross-references: GB:M62397; NID:G181034; PIDN:AAA52069.1; PID:G181035
C.Genetics:
A.Gene: GDB:MCC
A.Cross-references: GDB:128163; OMIM:159350
A.Map position: 5q21-5q21

```

```

Query Match 4.1%; Score 110; DB 2; Length 829;
Best Local Similarity 18.4%; Pred. No. 3.8; Matches 198; Indels 130; Gaps 19;
Matches 92; Conservative 81; Mismatches 198; Indels 130; Gaps 19;
QY 81 PSPGEKTI---QVPEAKLAGFLRYISMQNLAVIFDLDSYRTAREFDTSFGLKCLK 136
DB 307 PSTGELSTSSSSNDPIAKIARVK-----LSKTRGSSSSDRPVLG 349
QY 137 KVSIGGANLYRQAMSF-----NIFYHALVCAVLTNOETITAEQVKVLFED 185
DB 350 EISSIGVSSVAEHLAHSIQDCSNIOEIFQTLYSHG-----SAISESIRE-REV 398
QY 186 DERSTDSQQCSSEDEDIFEEETAQVSPRGKEKQWRAMPPLSLVQPVSNADWVLVRL 245
DB 399 ETERANSTIEHKSQNDLITLLE-----ECKNAERMMLVKYKSNATLALALQY 451
QY 246 HKLCMELCNNTYQMLHLDLENCEPPIIFKGPFFILPFSQSESSTPSTGFGSGKETPSE 304
DB 452 SEQCIIEAYELL-----ALAESEQSLILQFRAAGVSSPGDQSGDENITQM 498
QY 305 ----DDRQSQREHMGESLSIKAGG--GDLILPPSPKVEKKDPSRKKEWENAGNKITYMA 358
DB 499 LKRAHDCRKTENAKALMLKLDGSCGAFVAAGSVOP-----WESLSNSHTST 549
QY 359 ADKTSKIMTEYKRRQOHNLSAFPKEVVEKKG-----EPLG-----PRGD 401
DB 550 TSSTSSCDTEFTKDEQ-RLKDYIQQLKNDAAVKLTMLIELSHIDPLSDVYKPRGDS 608
QY 402 SPL-LQRP---QHLMDQGMHRSFAGPELLRQDKRPSGSGTSSLSVSVADAQIOAW 457
DB 609 ORLDLENVLMQELMAKEMVNAELKAQYLLEKEKK-----ALDKLSTRDAQDAY 660

```

```

QY 458 TMMVLTINQIILPDQTFALQ-----PAVFPCIS-----QLTCHVDIR 498
DB 661 LVHIEHLKSEVEEKEQRMRSLSSTSSGSKDKPKGECADASPALSIAELRTTSENELA 720
QY 499 VR--QAVREMLGRVGYDII 517
DB 721 AEFNALIREKKLAKARVOELV 741

```

RESULT 8

```

T24103
hypoetical protein R102.7 - Caenorhabditis elegans
C.Species: Caenorhabditis elegans
C.Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C.Accession: T24103
R.Berks, M.
submitted to the EMBL Data Library, March 1996
A.Reference number: Z19841
A.Accession: T24103
A.Status: preliminary; translated from GB/EMBL/DDBJ
A.Molecule type: DNA
A.Residues: 1-587 <WIL>
A.Cross-references: EMBL:Z70309; PIDN:CAA94361.1; GSPDB:GN00022; CESP:R102.7
A.Experimental source: clone R102
C.Genetics:
A.Gene: CESP:R102.7
A.Map position: 4
A.Introns: 19/2; 111/3; 147/1; 270/2; 407/1; 481/3; 529/2

```

```

Query Match 4.1%; Score 109.5; DB 2; Length 587;
Best Local Similarity 20.7%; Pred. No. 2.5; Matches 132; Indels 129; Gaps 19;
Matches 84; Conservative 60; Mismatches 132; Indels 129; Gaps 19;

```

```

QY 77 FVKG---PSPGEKTIQVPEAKLAGFLRYISMQNLAVIFDLDSYRTA--REPDSPGL 131
DB 6 FRKGVQPISSRKRTTLQNLKKARFVRLVFRVG-----FLDRHDTCHEREVQLTEDD 60
QY 132 KCLIKVSGIGGANLYRQAMSFNIFYHALVCAVLTNOETITAEQVKVLFEDDERSTD 191
DB 61 QKLIKLVK--EGETNNIKQSGKLFDDL--TWSCGSLI-----YKRGATITETNFNE 107
QY 192 SSQCSSEDEDIFETNOVSPR-----GKEKQWRAMPPLSLVQPVSNADWVLVRLH 246
DB 108 TMDPFDNEAAVLDVLQKRSKEKEKKKXEMDALDELTHVTSNDRSLSLR-- 165
QY 247 KLCMELCNNTYQMLHLDLENCEPPIIFKGPFFILPFSQSES-----TPSTGFGSGKET 301
DB 166 -----IIPDLHNSQSRKGLKLAPOST--IREHT 189
QY 302 PS-----EDDRSQREHMGESLSIKAGGDLILPPSPKVEKKDPSRKKEWENAGNKITY 356
DB 190 VSVOPTLEDMDTQK-----VKKPSSRPKHK-PSTIRE-----HT 224
QY 357 MAAKTISKLMTE---YKKRQOHNLSAFPKEVVEKK-----GEPLG 396
DB 225 LSVQKTVEDMDTQKVVKKSSRPKHKPSTMEVLKKKKSAATISKETCOLHTKFGCLFG 284
QY 397 -PRGQDSP-----LLQRPQHMDGQMRHSFAGPELLRQDKR 433
DB 285 TVTEPNNSPYQNLIDLSPRSKVDNPMQFLPSPINPKATEIRK 329

```

RESULT 9

```

A1948
alpha-fetoprotein enhancer-binding protein - human
N.Alternate names: ATBF1 protein
C.Species: Homo sapiens (man)
C.Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 15-Oct-1999
C.Accession: A1948
R.Moringa, T.; Yasuda, H.; Hashimoto, T.; Higashio, K.; Tamaoki, T.
Mol. Cell. Biol. 11, 6041-6049, 1991
A.Title: A human alpha-fetoprotein enhancer-binding protein, ATBF1, contains four homeod

```



```

QY 219 RQWRMRMP-----LSQVFNADMMVLVCKLHKLQMLCUNNYIQHMLDLCKMEPPIFKG 275
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 247 -SFAAGIPLNNKPAISSSSSNPKFVQLMR-----TIAPVSTSS 284
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 276 DPFLLPFSQSSSTPSTGSGKETPSEDDNRQSQSEHMGESLSLAKGDDLPP--- 311
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 285 RP--DIPFPVTSAP-----RPNQ-----RPLPHNAPKLEPVK-----VPIPSLARK 324
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 332 -----SPKVEKKDPSRK--KEMENAGNKIYTAADKTIKMLTEYKGRQOHNLSA 381
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 325 INSSAVPLRDRPIELKDPKALKD--SNIVLKGITIKPRFVIEFPKPPFKKPKQBEQ--- 379
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 382 FPKVEYKVEKKGP LG-PRGQDSPLLQRP----- 408
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 380 --EHTNGEKKPILDEPIIIDEPRTIRPBEHCALFNNLSTPKPRHGENTTSIRTVIMPK 457
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 409 -QHLMDQGMRRHSFASGPELLRQDKRPS 436
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 438 SMHLKNGE-----SNGPSLFATNKRPSP 461
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 12

Vitellogenin 1 precursor [validated] - chicken
N:Contans: 11povtellin I; 11povtellin II; phosvitin; yolk glycoprotein 42k
C:Species: Gallus gallus (chicken)
C:Date: 15-Oct-1993 #sequence_revision 15-Oct-1999 #text_change 18-Aug-2000
C:Accession: T730088, S55681
R:Mauchli, N.; Yamamura, J.; Adachi, T.; Aoki, N.; Nakamura, R.; Matsuda, T.
submitted to the EMBL Data Library, November 1996
A:Description: cDNA cloning and estrogen-induced expression of chicken vitellogenin I
#:Reference number: Z20557

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1912 <MBL>
A:Cross-references: EMBL:D89547; PIDD:BAI3973.1
A:Experimental source: liver
R:Yamamura, J., Adachi, T., Aoki, N., Nakajima, H., Nakamura, R., Matsuda, T.
Biochim. Biophys. Acta 1244, 384-394, 1995
A>Title: Precursor-product relationship between chicken vitellogenin and the yolk protein
A:Reference number: S55680; MUID:95322425; PMID:7599159
A:Accession: S55681
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1628-1639 <YAM>
C:Superfamily: vitellogenin
C:Keywords: egg yolk; glycoprotein; liver; phosphoprotein
F:1-15/Domain: signal sequence #status predicted <Sig>
F:16-1912/Product: vitellogenin I #status predicted <Mat>

Query Match 4.0%; Score 108; DB 2; Length 1912;
 Best Local Similarity 20.3%; Pred. No. 18;
 Matches 73; Conservative 39; Indels 137; Gaps 10

```

QY 176 EOVKKVLEJEDDERSTSSQOCSEDEDIIFEETAAQVSPRGEXKROWRAMPLLSVQPVSN 235
QY 1076 KRVKKILDDTDNOAARNRNRSSSSASSTISESESTTSPPSSDDNRAS----- 112
Db 236 ADWVWLVRKHLKCMELCNNTIYQHLLDLENCMEEPPIFKDPPFLPFSOSE----- 287
QY 1125 -----QGPQIMAKRQSAANEKEYP 1146
Db 288 -----SSTPSTGFGFSKTEPSEDRDSOSREHNGESTLSKAGGDLLEPPS 332
QY 1147 FGDSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSKXSSS---RS 1207
Db 333 PKVEKQDPSRKKEWENAGNKLYTTAAADKTI SKLMTTEYKRRQOAHLSAFPKEVKEVKKG 392
QY 1203 SKSNRSSSSSSSKSSSSSSSKSNKSGSSSSSSSKAGCTQKXKAKQKTTSPFPAASAAE--G 1266
Db 393 EPLRGQDSEPLRQPHILMDQGMKRRHSFSAQPELLRQDKRPRGSGTGS---SLSVSVD 449

```

Db 1261 E-----RSYH--EQKQETSSSSSSSRASNSRSTSSSTSSSESGVSHRQ 1305
 Qy 450 AEAQIQAATNNVLTVLNQIQI--LPDQTFALQPAVFPCCISQLTCHVTDIRVROAVREW 506
 Db 1306 WKQREAEATKRVKSGQPSHSHSYDIPNEMETYL-PKVY-----RLRFRSAHHW 1352

RESULT 13

peptide synthetase [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:date: 14-Dec-2001 #sequence revision 14-Dec-2001 #text change 09-Dec-2002
C:Accession: AB2136
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriyuchi, N.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, T.; Yamada, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-2617 <KUP>
A:Cross-references: GB:BA000019; PIDN:BAB74343.1; PID:g17131737; GSPDB:GNO0179
A:Experimental source: strain PCC 7120

C:Genetics:
A:Gene: a112644

C;Keywords: carrier protein, phosphopantetheine, phosphoprotein
C;Superfamily peptide synthetase psd; acetate-CoA ligase homology; acyl carrier protein
F;I025/Binding site: phosphopantetheine (Ser) (covalent) #status predicted
F;2068/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match	3.9%;	Score 106.5;	DB 2;	Length 2617;
Best Local Similarity	20.2%;	Pred. No. 36;		
Matches 75;	Conservative 49;	Mismatches 125;	Indels 123;	Gaps 16;

2 IRAMAQQVFM LDTQCS-----PKTPN-----NPDHA 27

Db 2288 IDGISWRVFLLEDQLAYQQQLSGGQIQLPKTTSYQQWANKLQEHWTWSADLQAAFNWYTS 234

QY 28 QSCOLIEIPDEKPNKHTKKSIFREIVSLSHOVLQ-----NLYDILIEEFVK 79

DB 2348 PIVQIIPPLPVDPAGSNIMADVDTISYFLSVIDIQNLLEQVPPAKYKQINDVLLTALVL 240

Db 2408 AFQWTGNNRLVELEGHREDIFPISINSRTWMTS----LFPVLIDYPSA--DL 2455

128 SPGLKCL---LKKVSGIGCANLYRQSAHSFNIFYHALVCAYLTNQETITAEQVKVLE 184

Db 2460 GISLKAIEQLRQIPDRGISYGLIRYLASP-----TIRDTIKALPIQVRF- 2505

185 DDERSTDSQCCSEDEDFEETLQAVSPR--GKEKQWRARMLLSQPV-----SNA 236

Db 2506-----NYLGQSDQIFSEBSLFTASESIGHSRSSRGKNTLLIENSLVYGKLRG 2555

Db 2556 DWTY-SKQLHRRQTATLAENYQILSLIHQCL-----TPRV 2592

QY 294 GGFSGKETPSED 305

Db 2593 SCPTSPDPQMD 2604

RESULT 14

B72855
global transactivator-like protein - Autographa californica nuclear polyhedrosis virus

C-Date: 12-Nov-1999 #sequence revision 12-Nov-1999 #text change 12-Nov-1999
A-Name: dsDNA virus
C-species: Autographa californica nuclear polyarthritis virus, ACNPV

C:\Accession: B72855

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 11:46:57 ; Search time 23.511 Seconds

(without alignments)
1036.105 Million cell updates/sec

Title: US-09-991-681-27

Perfect score: 2698
Sequence: 1 RIRAMAQOVEMLDTCSPKT.....VRQAVREMLGRVGRVDIV 518

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	126	4.7	2278	1	PAB1_YEAST
2	124	4.6	2150	1	SDC3_CAEEL
3	113	4.2	482	1	YGR2_CAEEL
4	110	4.1	829	1	CRCM_HUMAN
5	109.5	4.1	342	1	HXCA_HUMAN
6	108.5	4.0	872	1	S3B2_HUMAN
7	108.5	4.0	962	1	SOUR_DROME
8	108.5	4.0	3703	1	ABF1_HUMAN
9	108	4.0	1912	1	VTI1_CHICK
10	107	4.0	1762	1	DROQ_HUMAN
11	106	3.9	390	1	MEI1_XENLA
12	106	3.9	8545	1	ANCL_CAEEL
13	105	3.9	506	1	GTA_NFVAC
14	105	3.9	1214	1	BRF3_HUMAN
15	104.5	3.9	876	1	VP41_CAEEL
16	103.5	3.8	1302	1	GRPS_DROME
17	103	3.8	1502	1	MRP6_RAT
18	103	3.8	3726	1	ABF1_MOUSE
19	102	3.8	457	1	CMGA_HUMAN
20	101.5	3.8	4486	1	DTH9_HUMAN
21	100.5	3.7	453	1	SSF2_YEAST
22	100	3.7	1268	1	VEIN_HUMAN
23	99.5	3.7	342	1	HXCA_MOUSE
24	99.5	3.7	449	1	ENGA_MYCN
25	99.5	3.7	527	1	RG2_MOUSE
26	99.5	3.7	649	1	TOP3_SALTI
27	99.5	3.7	1534	1	RRE1_CANFA
28	99	3.7	456	1	YVEO_YEAST
29	99	3.7	1207	1	DML1_ARATH
30	99	3.7	1785	1	BIG2_HUMAN
31	98.5	3.7	449	1	CMGA_BOVIN
32	98.5	3.7	962	1	L436_CAEEL
33	98.5	3.7	986	1	DPOL_NPVBM

34	98.5	3.7	1341	1	ACIN_HUMAN	O9ukv3 homo sapien
35	98	3.6	1171	1	Y637_HUMAN	O75132 homo sapien
36	97.5	3.6	743	1	MYBB_XENLA	P52551 xenopus lae
37	97.5	3.6	2314	1	AKA6_RAT	O9wvc7 ratius norv
38	97	3.6	477	1	MEI2_HUMAN	O14770 homo sapien
39	97	3.6	700	1	TRDN_CANFA	P82179 canis fami
40	97	3.6	870	1	Y563_HUMAN	O60309 homo sapien
41	97	3.6	1369	1	MSH5_CAEEL	O19272 caenorhabdi
42	96.5	3.6	693	1	SC13_ARATH	O9F019 arabidopsis
43	96.5	3.6	906	1	I4C2_MOUSE	O62448 mus musculu
44	96.5	3.6	2195	1	SC16_YEAST	P48415 saccharomyc
45	96.5	3.6	2312	1	CCAE_HUMAN	Q15878 homo sapien

ALIGNMENTS

RESULT 1
FAB1_YEAST
ID FAB1_YEAST STANDARD; PRT; 2278 AA.
AC P34756;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 1-phosphatidylinositol-3-phosphate 5-kinase FAB1 (EC 2.7.1.150)
DE (Phosphatidylinositol-3-phosphate 5-kinase) (Type III PIP kinase).
GN FAB1 OR YFR019W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_Taxid=4932;
RN [1]
RP MEDLINE=95392039; PubMed=7663021;
RA Yamamoto A., Dewald D.B., Boronenkov I.V., Anderson R.A., Emr S.D.,
RA Koshland D.;
RT "Novel PI (4)P 5-kinase homologue, Fab1p, essential for normal vacuole
RT function and morphology in yeast."
RL Mol. Biol. Cell 6:525-539(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=95400292; PubMed=7670463;
RA Murakami Y., Natou M., Hagiwara H., Shibata T., Ozawa M.,
RA Sasaana S.-I., Sasaana M., Tsuchiya Y., Soeda E., Yokoyama K.,
RA Yamazaki M., Tashiro H., Eki T.;
RT "Analysis of the nucleotide sequence of chromosome VI from
RT Saccharomyces cerevisiae."
RL Nat. Genet. 10:261-268(1995).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=99035548; PubMed=9811604;
RA Cooke F.T., Dove S.K., McEwen R.K., Painter G., Holmes A.B.,
RA Hall M.N., Mitchell R.H., Parker P.J.;
RT "The stress-activated phosphatidylinositol 3-phosphate 5-kinase Fab1p
RT is essential for vacuole function in S. cerevisiae."
RL Curr. Biol. 8:1219-1222(1998).
RN [4]
RP SIMILARITY TO TCPI/CTP FAMILY.
RX MEDLINE=95314774; PubMed=7794526;
RA Waldmann T., Lupas A., Kellermann J., Peters J., Baumeister W.;
RT "Primary structure of the thermosome from Thermoplasma acidophilum."
RL Biol. Chem. Hoppe-Seyler 376:119-126(1995).
-1- FUNCTION: Catalyzes the phosphorylation of phosphatidylinositol-3-
phosphate on the fifth hydroxyl of the myo-inositol ring, to form
phosphatidylinositol-4,5-bisphosphate. Required for endocytic-
vacuolar pathway and nuclear migration. The product of the
reaction it catalyzes functions as an important regulator of
vacuole homeostasis perhaps by controlling membrane flux to and/or
from the vacuole.
-1- CATALYTIC ACTIVITY: ATP + 1-phosphatidyl-1D-myo-inositol 4-
monophosphate = ADP + 1-phosphatidyl-1D-myo-inositol 4,5-
bisphosphate.


```

Db 1327 TRGSPKVKYENLKRRIOQRQCKEBAHEPE-VVEEQOVEPEVPEVHEPVAPAAQL 1385
Qy 229 SVQPVSNADWWVLVRLHLKLCMELCNNTYIQMHLDLNCEMP-----PIFGDPFFILPS 283
Db 1386 ETEPI-----EQIEEPDKVFEIITA-----LPL 1410
Qy 284 FOSSSSTPSTGFGSGKETPSEDD-----RSQSRHEWG 315
Db 1411 FETSPVPAPAEGNIPFRAHSDDDVQVISSETDPNGPINLVEQVQNDKLTAYQSTEEILG 1470
Qy 316 ESLSLKAGCG-----DLLPSPKVKCKDPSRKKEWEN--AGNKITYMAA 359
Db 1471 EYGEIDEXAGAPSEIIVHDEVLODEVLPNPKSKKGRRRKTPPHIAKAKVFT--- 1527
Qy 360 DKTISKLMTE-----YKKRKOQNLISA-----FPKVEYKEKKGEP-----LGPR 398
Db 1528 --SISK--TEIELAPPTQOSRKMANVSSBEATATRQRAKVEENDSDVSRVLTPE 1583
Qy 399 GQDSPLLQRPQHLMDQGMHRSFSAGPELLRQDKRPRSGTSSLSVS----- 446
Db 1584 PEDLHETRPGHVGEGETPSTLRTGRE-----STASSVKTISRSLKFLSKNKP 1632
Qy 447 VRDAEQIQAWNTN 459
Db 1633 VPRMRIOQAGTN 1645

```

RESULT 3

```

YSR2_CAEEL STANDARD; PRT; 482 AA.
ID YSR2_CAEEL
AC Q09950;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 54.6 kDa protein F59B10.2 in chromosome II.
CN F59B10.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditidae; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Lloyd C.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; Z48716; CA88601.1; -
DR PIR; T22981; T22981.
DR WormPep; F59B10.2; CE03591.
KM Hypothetical protein.
FT DOMAIN 214 282 SER-RICH.
FT DOMAIN 228 241 POLY-SER.
FT DOMAIN 276 282 POLY-SER.
FT DOMAIN 378 384 POLY-SER.
FT DOMAIN 397 412 ARG-RICH.
FT DOMAIN 404 412 POLY-ARG.
SQ SEQUENCE 482 AA; 54641 MW; 1BD7D31B9DDAA74B CRC64;

```

Query Match 4.2%; Score 113; DB 1; Length 482;

Best Local Similarity 21.9%; Pred. No. 1;

Matches 85; Conservative 46; Mismatches 128; Indels 130; Gaps 16;

Qy 118 SYRRAREDDTS-----PGLKCLIKVSGIGGAAANLYRSANSFNITYFALVCAVITN 169

Db 81 SLKDIRLRNSIDQSLIKPPELKMRLSRVAG-----EESKRHAHFESIOEEKIS 130

```

Qy 170 QETITAEQVKVLFEDDERSTDSQCCSEDEDIIEETAQ----VSPRGKEKQWRAMP 226
Db 131 EDLPQAQLPKSPF---KKAIQSGSESDSDSIIFDFVFEVLPSPRR----- 175
Qy 227 LLSVQPVSNADWWVLVRLHLKLCMELCNNTYIQMHLDLNCEMPPI-----FKGDPF 278
Db 176 --KPAPARTAPIV-----VEKKIEKPAVEQORAKKKEKPT 210
Qy 279 FILSPQ-----SSSTPSTGFGSGKETPSEDD-----SQSRHEWGSLSLKAGGD 326
Db 211 PTSESSFESSDSSSTSSSTSSSESSSASSESSKSESOVSSKSTSKASSKAQSD 270
Qy 327 LLLPSPKVKCKDPSRKKEWENAGNKITYMAADKITSKL---MTEYKKRKOQNLISAF 382
Db 271 F-----ESEKSSS-----SASTISKYTPKLDLPQTKKPKDKRAK 307
Qy 383 PKEVYKEKKGEPILQPRGQDSPLLQRPQHLMDQGMHRSFSAGPELL----- 428
Db 308 PDDIRQKKEPI-PYEDFIPKLSRSNSSEKSTVRETNTLESLKTKLKINSSEKME 366
Qy 429 --RODKR--PRSGTSSLSVVRDAEQ 453
Db 367 KPRDIRRAPRSSSSSSS---TLRDARE 392

```

RESULT 4

```

CRCM_HUMAN STANDARD; PRT; 829 AA.
ID CRM_HUMAN
AC P23508;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Colorectal mutant cancer protein (MCC protein).
CN MCC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=9116485; PubMed=1848370;
RA Kinzler K.W., Nilbert M.C., Vogelstein B., Bryan T.M., Levy D.B.,
RA Smith K.J., Preisinger A.C., Hamilton S.R., Hedge P., Markham A.,
RA Carlson M., Joslyn G., Groden J., White R., Miki Y., Miyoshi Y.,
RA Nishihio I., Nakamura Y.;
RT "Identification of a gene located at chromosome 5q21 that is mutated
RT in colorectal cancers";
RL Science 251:1366-1370(1991).
RN [2]
RP VARIANTS COLORECTAL CANCER MUTANTS.
RX MEDLINE=9135211; PubMed=1651563;
RA Nishihio I., Nakamura Y., Miyoshi Y., Miki Y., Ando H., Horii A.,
RA Koyama K., Utsunomiya J., Baba S., Hedge P., Markham A., Krush A.J.,
RA Petersen G., Hamilton S.R., Nilbert M.C., Levy D.B., Bryan T.M.,
RA Preisinger A.C., Smith K.J., Su L.-K., Kinzler K.W., Vogelstein B.;
RT "Mutations of chromosome 5q21 genes in FAP and colorectal cancer
RT patients.";
RL Science 253:665-669(1991).
CC -I- FUNCTION: CANDIDATE FOR THE PUTATIVE COLORECTAL TUMOR SUPPRESSOR
CC GENE LOCATED AT 5021.
CC -I- TISSUE SPECIFICITY: EXPRESSED IN A VARIETY OF TISSUES TYPES.
CC -I- DISEASE: IS PROBABLY INVOLVED IN EARLY STAGES OF COLORECTAL
CC NEOPLASIA IN BOTH SPORADIC AND FAMILIAL TUMORS.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC -----
DR EMBL; M62397; AA552069.1; -
DR PIR; A38434; A33166.
DR Genew; HGNC:6935; MCC.
DR MIM; 159350; -.
DR GO; GO:0004872; F:receptor activity; TAS.
DR GO; GO:0008181; F:tumor suppressor; TAS.
DR GO; GO:0007165; P:signal transduction; TAS.
KM Anti-oncogene; Disease mutation.
FT VARIANT 267
FT VARIANT 267
FT VARIANT 486
FT VARIANT 486
FT VARIANT 490
FT VARIANT 490
FT VARIANT 506
FT VARIANT 506
FT VARIANT 698
FT VARIANT 698
SQ SEQUENCE 829 AA, 93055 MW, 5D45225FD91CA18F CRC64;

Query Match 4.1%; Score 110; DB 1; Length 829;
Best Local Similarity 18.4%; Pred. No. 3.5;
Matches 92; Conservative 81; Mismatches 198; Indels 130; Gaps 19;

QY 81 PSRGEKTI---QVPEAKLGFLLRYISQNLAVFDLLDLYRTAREFDTSFGLKLIK 136
DB 307 PSTGLSTSSSSNDPIAKIAERVK-----LSKTRBSSSSDRPVGS 349
QY 137 KVSIGIGANALYRQAMSF-----NIYFALVCAVLTNOETITAEQVKVLPED 185
DB 350 EISSIGVSSVAEHLAHSIQDCSNIOEIFQTLYSHG-----SAISEKIRE--FEV 398
QY 186 DERSTDSQQSCSEDEDIFEETAQVSPPRGKEKQWRAMPILSVQVSNADWVLVRL 245
DB 399 ETERNSRIEHLKSNQDLITLLE-----ECKSNAERMSWLVGYESNATRLALQY 451
QY 246 HKLMELCNNYIQMLDLNLCMEBPPIKGFPIFLPSQSSSTPFGSGEKTPSE- 304
DB 452 SEQCIENAYELD-----ALMBESQSLTIGQRAAGVSSSPDQSGDENTITQ 498
QY 305 ----DRSQSREHMESSLSLKAG--GDILLPPSPKVEKQPSRKKEWMENAGNKIYTMA 358
DB 499 LKRAHDCRKTAKENAKALMLKLDGSCGAFNAGSVQ-----WESLSSSHST 549
QY 359 ADKTIISKMTYEKKKQKHNLAPPEKAVKEKG-----EPLG---PRQD 401
DB 550 TSSSTASCDTEFTKEDEQ-RLKDYIQQLKNDRAAVKLTWLELESIHIDPLSYDVKPRDS 608
QY 402 SPL-LQRP---QMLDQGMRSFSPAGBELLRQDRPSGSGSSLSVSDAEOIQAW 457
DB 609 QRLDENAVLMQELAMKEEMALKAQIYLLKEKK-----ALEKLSSTREQOQAY 660
QY 458 TNNVLTVINOQILPDQTFALQ-----PAVPCIS---QLTCHVTDIR 498
DB 661 LVHIEHLSVYEQEQRMRSLSSSTSSGSKKPKGECADASPALSLAEIRTCSENELA 720
QY 499 VR--QAVREWLGRVGRVYDI 517
DB 721 AEFTNAIRREKKLKARVOELV 741

RESULT 5
HXCA HUMAN STANDARD; PRT; 342 AA.
AC Q9NYD6; O15219; O15220; Q9BYD5;
DT 16-OCT-2001 (Rel. 40, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DS Homeobox protein Hox-C10.
GN HOXC10.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OK NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20296799; PubMed=10835276;
RA de Stanchina E., Gabbellini D., Norio P., Giacca M., Peverali F.A.,
RA Riva S., Falschi A., Biamonti G.,
RT "Selection of homeotic proteins for binding to a human DNA replication
RT origin."
RT J. Mol. Biol. 299:667-680(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Cervix;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan J., Moore T., Max S.I., Wang J., Heisen F.,
RA Ditschenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulys S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalley D.E.,
RA Schermer A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE OF 94-106 AND 258-297 FROM N.A.
RX MEDLINE=98019011; PubMed=9357979;
RA Flagiello D., Gibaud A., Dutrillaux B., Pouyon M.F., Malfoy B.;
RT "Distinct patterns of all-trans retinoic acid dependent expression of
RT HOXB and HOXC homeogenes in human embryonal and small-cell lung
RT carcinoma cell lines."
RT FEBS Lett. 415:263-267(1997).
CC -1- FUNCTION: SEQUENCE-SPECIFIC TRANSCRIPTION FACTOR WHICH IS PART OF
CC A DEVELOPMENTAL REGULATORY SYSTEM THAT PROVIDES CELLS WITH
CC SPECIFIC POSITIONAL IDENTITIES ON THE ANTERIOR-POSTERIOR AXIS.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: BELONGS TO THE ABD-B HOMEBOX FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; AF255675; AA67759.1; -
DR EMBL; BC001293; AAH01293.1; -
DR EMBL; X99684; CAA6799.1; -
DR EMBL; X99685; CAA6800.1; -
DR HSSP; P02833; 9ANT.
DR TRANSFAC; T03328; -
DR Genew; HGNC:5122; HOXC10.
DR MIM; 605560; -.
DR GO; GO:0003702; F:RNA polymerase II transcription factor acti. .; TAS.
DR GO; GO:0008284; P:positive regulation of cell proliferation; TAS.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00045; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.

```

KM Homeobox; DNA-binding; Developmental protein; Nuclear protein;
 KW Transcription regulation.
 FT DNA BIND 268 327 HOMEBOX.
 FT CONFLICT 118 118 K -> N (IN REF. 1).
 FT CONFLICT 265 265 A -> G (IN REF. 3).
 FT CONFLICT 271 271 MISSING (IN REF. 3).
 SQ SEQUENCE 342 AA; 38072 MW; BD8127FD43C2A37B CRC64;
 Query Match 4.1%; Score 109.5; DB 1; Length 342;
 Best Local Similarity 19.3%; Pred. No. 1.2;
 Matches 64; Conservative 38; Mismatches 108; Indels 121; Gaps 11;
 QY 144 AANLYROSAMSFNITFHALVCAVLTNOETITAEQVKVLPEDDERSTDSQQCCSSEDEI 203
 DB 28 SAGWMOGSDPFI-----CGVM-----RGCLAPLSIKRDEGS 60
 QY 204 FEETAAQVSPR-----GKEKQWARMPLLSVQPSNADWWLVKRLKLMELCNN 255
 DB 61 SPSLALNTYPSYLSQLDSWGPDKAAYRLQPV--GRPLSSGCSYPSVKEENVCCMYSAEK 118
 QY 256 -----YIQMHLDLNCEMEEPPIFKGDPFPIPSFQSESTPSTGSGKETPSEDDR 307
 DB 119 RAKSGPEALYSHLPESCLEGEHPVPSYRASPYSALDKTPHCSGANDPEAFEGRA 178
 QY 308 SQS---REHMGESLKLKAGGDLPLPSPKVEKQPSRKKEWENAGNKITYMAADKITSK 365
 DB 179 SLNPRAEHL--ESPOL---GKAVSPETPKSDSQTPS----- 210
 QY 366 LMTVEYKKKQQAHLNAPFKVEYKVEKGPPLPGRQDPSLPLRPHLMDQGMKHSFSGAP 425
 DB 211 -----PNEIKTQD--SLAAPKSGPS----- 228
 QY 426 ELLRQDKPRSGTSSLSVSVRDAEQIOA 456
 DB 229 -----ESEKERAKAADSPDTSDNKAKEIKYA 255
 RESULT 6
 S3B2 HUMAN
 ID_S3B2 HUMAN STANDARD; PRT; 872 AA.
 AC 013435.
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Splicing factor 3B subunit 2 (Spliceosome associated protein 145) (SAP 145) (SF3B150) (Pre-mRNA splicing factor SF3b 145 kDa subunit).
 GN SF3B2 OR SAP145.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 151-159 AND 794-817.
 RX MEDLINE=96154048; PubMed=8566756;
 RA Gozani O., Feld R., Reed R.;
 RT "Evidence that sequence-independent binding of highly conserved U2 snRNP proteins upstream of the branch site is required for assembly of spliceosomal complex A.";
 RL Genes Dev. 10:233-243(1996).
 RN [2]
 RP CHARACTERIZATION OF THE SPLICOSOME.
 RX MEDLINE=20337962; PubMed=10882114;
 RA Das R., Zhou Z., Reed R.;
 RT "Functional association of U2 snRNP with the ATP-independent spliceosomal complex E.";
 RL Mol. Cell 5:779-787(2000).
 CC -1- FUNCTION: SUBUNIT OF THE SPLICING FACTOR SF3B REQUIRED FOR 'A' COMPLEX ASSEMBLY FORMED BY THE STABLE BINDING OF U2 SNRNP TO THE BRANCHPOINT SEQUENCE (BPS) IN PRE-MRNA. SEQUENCE INDEPENDENT BINDING OF SF3A/SF3B COMPLEX UPSTREAM OF THE BRANCH SITE IS ESSENTIAL. IT MAY ANCHOR U2 SNRNP TO THE PRE-MRNA. MAY ALSO BE INVOLVED IN THE ASSEMBLY OF THE 'E' COMPLEX. BELONGS ALSO TO THE MINOR U12-DEPENDENT SPLICOSOME, WHICH IS INVOLVED IN THE SPLICING

CC OF RARE CLASS OF NUCLEAR PRE-MRNA INTRON.
 CC -1- SUBUNIT: COMPONENT OF SPLICING FACTOR SF3B WHICH IS COMPOSED OF FOUR SUBUNITS: SF3B4/SAP49, SF3B3/SAP130, SF3B2/SAP145, SF3B1/SAP155. SF3B ASSOCIATES WITH THE SPLICING FACTOR SF3A AND A 12S RNA UNIT TO FORM THE U2 SMALL NUCLEAR RIBONUCLEOPROTEIN COMPLEX (U2 SNRNP). SF3B2 INTERACTS DIRECTLY WITH SF3B4.
 CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -1- SIMILARITY: TO YEAST CUS1. SOME, TO C. ELEGANS ZK632.11.
 CC -1- SIMILARITY: Contains 1 SAP domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.ebi.ch/announce/> or send an email to license@ebi.ebi.ch).
 CC -----
 CC EMBL; U41371; AAA97461.1; -.
 CC Genew; HGNC:10769; SF3B2.
 CC DR GK; O13435; -.
 CC DR MIM; 605591; -.
 CC DR GO; GO:0005681; C:spliceosome complex; TAS.
 CC DR GO; GO:0008248; F:Pre-mRNA splicing factor activity; TAS.
 CC DR GO; GO:0006371; P:mRNA splicing; TAS.
 CC DR InterPro; IPR006568; PSP.
 CC DR InterPro; IPR003034; SAP.
 CC DR Pfam; PF04037; DUF382; 1.
 CC DR Pfam; PF04046; PSP; 1.
 CC DR Pfam; PF02037; SAP; 1.
 CC DR SMART; SM00581; PSP; 1.
 CC DR SMART; SM00513; SAP; 1.
 CC DR PROSITE; PS50800; SAP; 1.
 CC KW Spliceosome; mRNA processing; mRNA splicing; Nuclear protein.
 FT DOMAIN 1 35
 FT DOMAIN 68 73
 FT DOMAIN 81 90
 FT DOMAIN 106 109
 FT DOMAIN 226 230
 FT DOMAIN 269 274
 FT DOMAIN 308 312
 FT DOMAIN 676 679
 FT DOMAIN 697 703
 SQ SEQUENCE 872 AA; 97656 MW; AED6E9FD0DA5DE31 CRC64;
 Query Match 4.0%; Score 108.5; DB 1; Length 872;
 Best Local Similarity 19.4%; Pred. No. 4.7; Indels 97; Gaps 13;
 Matches 63; Conservative 46; Mismatches 118;
 QY 168 TNOETITAEQVKVLPEDDERSTDSQQCCSSEDEIPEETAQVSP-----RGKEKR 219
 DB 259 SRQEKMSQO-----EEEMETDARSSLSGQASSETEDIVSYSKKKRKRKKKKK 312
 QY 220 QWRAMPPLLSVQPSNADWWLVKRLKLMELCNNYIQMHLNCEMEEPPIFKGDPF 279
 DB 313 PORVR-----CVSSSSSD-----REKDSSTRSGSDPADVELEYTEPEPEIYEPNFI 362
 QY 280 ILPSFQS-----ESSPTSGFSGKETPSSDDDSQSEHNGES 317
 DB 363 FKRIFEAFKLTLDVYKKEKEPEKLDKLXENSAAPKKGKFEDEHKSDSDSDDEQEK-- 419
 QY 318 LSLKAGGDDLPLPSPKVEKQPSRKKEWENAGNKITYMAADKITSKLTVEYKKRQKH 377
 DB 420 -----KBPATLSKKKLRMR-----FTYAE-----LKQLVARDVYEMH 455
 QY 378 NLSA-FPK---EYKVEKGEPL-----GPRGQDSPLLQRPQHLMDQ--QMRH 419
 DB 456 DVTAQDDKLVLVHKATRNSPVPRHWCFKKRYLOGKKGIKPPFELDPFKRTGIQMR 515
 QY 420 SFSAGPE-----LLRQDKPRSG 437
 DB 516 ALOEKEOKTKMKSIMREKVRPKMG 539

ID	SUR	DROME	STANDARD	PRT	962 AA.
AC	15-SEP-2003	(Rel. 42, Created)			
DT	15-SEP-2003	(Rel. 42, Last sequence update)			
DT	15-SEP-2003	(Rel. 42, Last annotation update)			
DE	Suppressor of underreplication protein.				
GN	SUR OR SU (UR) ES OR CG7859				
OS	Drosophila melanogaster (fruit fly).				
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;				
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;				
OC	Ephydroidea; Drosophilidae; Drosophila.				
OX	NCBI_Taxid=7227;				
RP	SEQUENCE FROM N.A., FUNCTION, SUBCELLULAR LOCATION, TISSUE				
RP	SPECIFICITY, AND DEVELOPMENTAL STAGE.				
RC	TISSUE=Embryo;				
RC	MEDLINE=21898319; PubMed=11901119;				
RC	Matunin I.V., Volkova E.I., Belyaeva E.S., Nabirochkina E.N.,				
RA	Pirokta V., Zhitulev I.F.;				
RA	"The Drosophila suppressor of underreplication protein binds to				
RT	late-replicating regions of polytene chromosomes."				
RL	Genetics 160:1023-1034(2002).				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=Berkley;				
RC	MEDLINE=20196005; PubMed=1073112;				
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,				
RA	Amaratunga P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,				
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,				
RA	Sutcliffe G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,				
RA	Brandon R.C., Rogers J.-H., Blazey R.G., Chamee M., Pfeiffer B.D.,				
RA	Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,				
RA	Abbil J.F., Agbayan A., An H.-O., Andrews-Pfannkoch C., Baldwin D.,				
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,				
RA	Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,				
RA	Botkova D., Botchan M.R., Bouck J., Brokstein P., Broctier P.,				
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,				
RA	Cherry J.M., Cui S., Dahle S., Dahle C., Davenport L.B., Davies P.,				
RA	de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,				
RA	Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,				
RA	Dubin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,				
RA	Fowler C., Gebrelian A.E., Garg N.S., Gelbart W.M., Glasser K.,				
RA	Glocke A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,				
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,				
RA	Hoslin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,				
RA	Jaitai W., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,				
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,				
RA	Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,				
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,				
RA	Meikulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,				
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,				
RA	Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pachet J.M.,				
RA	Palazzolo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,				
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,				
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,				
RA	Spieler E., Spreading A.C., Steplaton M., Strong R., Sun E.,				
RA	Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,				
RA	Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,				
RA	Williams S.M., Wozniak T., Worley K.C., Wu D., Yang S., Yao Q.A.,				
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,				
RA	Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,				
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,				
RT	"The genome sequence of Drosophila melanogaster."				
RL	Science 287:2185-2195(2000).				
RP	REVISIONS.				
RC	STRAIN=Berkley;				
RC	MEDLINE=22426069; PubMed=12537572;				
RA	Miera S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,				
RA	Hradecky P., Huang Y., Kamitani J.S., Milburn G.H., Prochink S.E.,				

Y	230	VQEVSNADNVWLVKRLHKLCHMLCNVNIQWMLDLENC-MEEPPIFKQDPFILLPSQSSS	288
D	288	IDNNPAQOKTKTLFETRLTSLSHSDVDAWSPLLFEVSSDDDE-----LT	333
O	170	QETITLQVQKVLFFEDDSDSTDSQQCCSDEDDIFEETLAQVSPRGKEKQMAPMLLS	229
D	237	---DFRLRRFR--HGFDSLPLVAPEQYKHWL-----NWLMAKNSQSSTISGSVCSTIAS	287
O	110	VIFDLILDSYRARRRFDSPGKILKLKKVSGISGAANLNYGQSMSPFIYHAIYCAVLIN	169
D	182	LYNVAVRLGRLHQYKQSPASFDRKPHLPDPKPEVFSKRIDLEEVYKQKGFSEYIK-----	236
O	69	LYDIL-----LEFPYKG-----PSPE--EKTIVQE-AKLAGEL-RYISMONIA	109
Q		Query Match 4.0%; Score 108.5; DB 1; Length 962;	
D		Best Local Similarity 21.1%; Pred. No. 5.4;	
		Matches 106; Conservative 65; Mismatches 178; Indels 153; Gaps 24	
CC		or send an email to license@lsb-8ib.ch .	
CC		EMBL; A02775592; CAB89187.1; -	
DR		EMBL; AE003546; AAF50110.3; -	
DR		FLYBase; FBgn0025355; SUUR.	
DR		InterPro; IPR000330; SNF2_N.	
DR		Pfam; PF00176; SNF2_N; 1.	
KR		SNF2_N; 1.	
Q		SEQUENCE 962 AA; 107624 MW; 5063FB8BE875C4F9 CRG64;	
CC		FUNCTION: Expressed throughout development. Weakly	
CC		expressed. Expressed at higher level in embryos and adult females.	
CC		-1- DEVELOPMENTAL STAGE: Expressed both maternally and zygotically.	
CC		-1- TISSUE SPECIFICITY: Expressed throughout development. Weakly	
CC		expressed. Expressed at higher level in embryos and adult females.	
CC		-1- DEVELOPMENTAL STAGE: Expressed both maternally and zygotically.	
CC		-1- TISSUE SPECIFICITY: Expressed throughout development. Weakly	
CC		expressed. Expressed at higher level in embryos and adult females.	
CC		-1- DEVELOPMENTAL STAGE: Expressed both maternally and zygotically.	
CC		-1- TISSUE SPECIFICITY: Expressed throughout development. Weakly	
CC		expressed. Expressed at higher level in embryos and adult females.	
CC		-1- DEVELOPMENTAL STAGE: Expressed both maternally and zygotically.	
CC		-1- TISSUE SPECIFICITY: Expressed throughout development. Weakly	
CC		expressed. Expressed at higher level in embryos and adult females.	
CC		-1- DEVELOPMENTAL STAGE: Expressed both maternally and zygotically.	
CC		-1- TISSUE SPECIFICITY: Expressed throughout development. Weakly	
CC		expressed. Expressed at higher level in embryos and adult females.	
CC		-1- DEVELOPMENTAL STAGE: Expressed both maternally and zygotically.	
CC		-1- TISSUE SPECIFICITY: Expressed throughout development. Weakly	
CC		expressed. Expressed at higher level in embryos and adult females.	
CC		-1- DEVELOPMENTAL STAGE: Expressed both maternally and zygotically.	
CC		-1- TISSUE SPECIFICITY: Expressed throughout development. Weakly	
CC		expressed. Expressed at higher level in embryos and adult females.	
CC		-1- DEVELOPMENTAL STAGE: Expressed both maternally and zygotically.	
CC		-1- TISSUE SPECIFICITY: Expressed throughout development. Weakly	
CC		expressed. Expressed at higher level in embryos and adult females.	
CC		-1- DEVELOPMENTAL STAGE: Expressed both maternally and zygotically.	
CC		-1- TISSUE SPECIFICITY: Expressed throughout development. Weakly	
CC		expressed. Expressed at higher level in embryos and adult females.	
CC		-1- DEVELOPMENTAL STAGE: Expressed both maternally and zygotically.	
CC		-1- TISSUE SPECIFICITY: Expressed throughout development. Weakly	
CC		expressed. Expressed at higher level in embryos and adult females.	
CC		-1- DEVELOPMENTAL STAGE: Expressed both maternally and zygotically.	
CC		-1- TISSUE SPECIFICITY: Expressed throughout development. Weakly	
CC		expressed. Expressed at higher level in embryos and adult females.	
CC		-1- DEVELOPMENTAL STAGE: Expressed both maternally and zygotically.	
CC		-1- TISSUE SPECIFICITY: Expressed throughout development. Weakly	
CC		expressed. Expressed at higher level in embryos and adult females.	
CC		-1- DEVELOPMENTAL STAGE: Expressed both maternally and zygotically.	
CC		-1- TISSUE SPECIFICITY: Expressed throughout development. Weakly	
CC		expressed. Expressed at higher level in embryos and adult females.	
CC		-1- DEVELOPMENTAL STAGE: Expressed both maternally and zygotically.	
CC		-1- TISSUE SPECIFICITY: Expressed throughout development. Weakly	
CC		expressed. Expressed at higher level in embryos and adult females.	
CC		-1- DEVELOPMENTAL STAGE: Expressed both maternally and zygotically.	
CC		-1- TISSUE SPECIFICITY: Expressed throughout development. Weakly	
CC		expressed. Expressed at higher level in embryos and adult females.	
CC		-1- DEVELOPMENTAL STAGE: Expressed both maternally and zygotically.	
CC		-1- TISSUE SPECIFICITY: Expressed throughout development. Weakly	
CC		expressed. Expressed at higher level in embryos and adult females.	
CC		-1- DEVELOPMENTAL STAGE: Expressed both maternally and zygotically.	
CC		-1- TISSUE SPECIFICITY: Expressed throughout development. Weakly	
CC		expressed. Expressed at higher level in embryos and adult females.	
CC		-1- DEVELOPMENTAL STAGE: Expressed both maternally and zygotically.	
CC		-1- TISSUE SPECIFICITY: Expressed throughout development. Weakly	
CC		expressed. Expressed at higher level in embryos and adult females.	
CC		-1- DEVELOPMENTAL STAGE: Expressed both maternally and zygotically.	
CC		-1- TISSUE SPECIFICITY: Expressed throughout development. Weakly	
CC		expressed. Expressed at higher level in embryos and adult females.	
CC		-1- DEVELOPMENTAL STAGE: Expressed both maternally and zygotically.	
CC		-1- TISSUE SPECIFICITY: Expressed throughout development. Weakly	
CC		expressed. Expressed at higher level in embryos and adult females.	
CC		-1- DEVELOPMENTAL STAGE: Expressed both maternally and zygotically.	
CC		-1- TISSUE SPECIFICITY: Expressed throughout development. Weakly	
CC		expressed. Expressed at higher level in embryos and adult females.	
CC		-1- DEVELOPMENTAL STAGE: Expressed both maternally and zygotically.	
CC		-1- TISSUE SPECIFICITY: Expressed throughout development. Weakly	
CC		expressed. Expressed at higher level in embryos and adult females.	
CC		-1- DEVELOPMENTAL STAGE: Expressed both maternally and zygotically.	
CC		-1- TISSUE SPECIFICITY: Expressed throughout development. Weakly	
CC		expressed. Expressed at higher level in embryos and adult females.	
CC		-1- DEVELOPMENTAL STAGE: Expressed both maternally and zygotically.	
CC		-1- TISSUE SPECIFICITY: Expressed throughout development. Weakly	
CC		expressed. Expressed at higher level in embryos and adult females.	
CC		-1- DEVELOPMENTAL STAGE: Expressed both maternally and zygotically.	
CC		-1- TISSUE SPECIFICITY: Expressed throughout development. Weakly	
CC		expressed. Expressed at higher level in embryos and adult females.	
CC		-1- DEVELOPMENTAL STAGE: Expressed both maternally and zygotically.	


```

FT DOMAIN 3227 3231 POLY-GLN.
FT DOMAIN 3376 3389 POLY-GLN.
FT DOMAIN 3392 3395 POLY-GLN.
FT DOMAIN 3507 3527 POLY-GLY.
FT DOMAIN 3597 3600 POLY-PRO.
FT DOMAIN 3636 3639 POLY-SER.
FT VANSPLC 1 914 Missing (in isoform B).
FT VARIANT 3374 3374 A->V.
FT VARIANT 3377 3384 /FTID=VAR_011694.
FT VARIANT 3527 3527 /FTID=VAR_011695.
FT CONFLICT 72 72 /FTID=VAR_011696.
FT CONFLICT 422 422 A->S (IN REF. 3).
FT CONFLICT 579 579 T->P (IN REF. 3).
FT CONFLICT 767 767 I->A (IN REF. 3).
FT CONFLICT 777 777 A->V (IN REF. 3).
FT CONFLICT 846 849 HHRV->RHUG (IN REF. 3).
FT CONFLICT 997 997 A->S (IN REF. 3).
FT CONFLICT 1150 1190 BEAIDVEGSETPADPEELADQEGASSQAERELTDS
-> GEMSHRGRPRGLGLVHLETSRGLLFEGDVTDPAGPH
VPY (IN REF. 3).
SQ SEQUENCE 3703 AA; 404468 MW; 0F62AF37D4DCFB56 CRC64;

Query Match 4.0%; Score 108.5; DB 1; Length 3703;
Best Local Similarity 20.6%; Pred. No. 33;
Matches 73; Conservative 45; Mismatches 91; Indels 145; Gaps 18;

OY 132 KCLKKVVGIGGAAALYRQASMFYIFHALCAVL-----NQEIIT 174
DB 1072 KHLQGHESGVE-----SCYHCVCANSTAKNLQHVSMKHQESLRK 1120
OY 175 AEQVKVLFEDDE-----RSTD-----SSQ 194
DB 1121 LQRLQKGLPEDEDLQGITTRCRPTDPEAIEDVEGSETPADPEELADQEGASS 1180
OY 195 QSSSEDEIFETPAQVSPRGKQKQARMPLSVQ-----VANA 236
DB 1181 QAEKLTSPATSKRISFPGSS-----PLSKRPKTAERKEPEQMYOCPYKYSNA 1233
OY 237 DMVNLVKRLH-----KLCMELCNN--YIOMHLD-----LENMEEPIFX 274
DB 1234 D-----VNLRYHAMTQHSVQPMRLCPDQDMNKNKHQLHLTHLSVAAPDCEVXLMTV 1289
OY 275 GDPPIILPSFQ-SESSPTSTGFGSGKETPSEDDRSQSEHMGESLSLAKG--GGDLLP 330
DB 1290 TTPENVMWSSMFLPAVDPDRGNSNLEBAGK--QETSDEDLCKNLPSASTHQSGLD--LK 1345
OY 331 PSPKVEKQDPSRKE-----WENNAKNIYMAADKTIKMTYK---KRKQ 375
DB 1346 PSP-----ADPGSVREDSGFICWKQCNQVF-----KTSAAALQTHNEVHAKRPQ 1390

RESULT 9
VITL_CHICK STANDARD; PRT; 1912 AA.
AC P87498;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Vitellinogenin I precursor (minor vitellinogenin) [Contains: Lipovitellin
I (LVI); Phosvitin (PV); Lipovitellin II (LVIi); YGP42].
GN VIT1 OR VIT1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
RX (1)
RY SEQUENCE FROM N.A.
RC TISSUE=Liver;

```

```

RA Mabuchi N., Yamamura J.-I., Adachi T., Aoki N., Nakamura R.,
RA Matsuda T.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE OF 1628-1639, AND IDENTIFICATION OF YGP42.
RX MEDLINE=95322425; PubMed=7599159;
RA Yamamura J.-I., Adachi T., Aoki N., Nakajima H., Nakamura R.,
RA Matsuda T.;
RT "Precursor-product relationship between chicken vitellinogenin and the
RT yolk proteins: the 40 kDa yolk plasma glycoprotein is derived from
RL the C-terminal cysteine-rich domain of vitellinogenin II."
RL Biochim. Biophys. Acta 1244:184-194(1995).
CC -1- FUNCTION: PRECURSOR OF THE EGG-YOLK PROTEINS THAT ARE
CC SOURCES OF NUTRIENTS DURING EARLY DEVELOPMENT OF OVIPAROUS
CC ORGANISMS.
CC -1- FUNCTION: PHOSVITIN IS BELIEVED TO BE OF IMPORTANCE IN
CC SEQUESTERING CALCIUM, IRON AND OTHER CATIONS FOR THE DEVELOPING
CC EMBRYO.
CC -1- TISSUE SPECIFICITY: PRODUCED BY THE LIVER, SECRETED INTO THE BLOOD
CC AND THEN SEQUESTERED BY RECEPTOR MEDIATED ENDOCYTOSIS INTO GROWING
CC OOCYTES, WHERE IT IS GENERALLY CLEAVED, GIVING RISE TO THE
CC RESPECTIVE YOLK COMPONENTS.
CC -1- INDUCTION: By steroids (estrogen).
CC -1- PMW: PHOSVITIN, AN EGG YOLK STORAGE PROTEIN, IS ONE OF THE MOST
CC HIGHLY PHOSPHORYLATED (104) PROTEINS IN NATURE.
CC -1- PMW: CATHEPSIN D IS RESPONSIBLE FOR INTRAOOCYTIC PROCESSING OF
CC VITELLOGENIN.
CC -1- PMW: MAY CONTAIN INTRAMOLECULAR DISULFIDE BONDS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
DR EMBL; D89547; BAAL3973.1; -.
DR PIR; T29088; T29088.
DR InterPro; IPR001747; Lipid transprt_N.
DR InterPro; IPR001846; VWF D.
DR Pfam; PF01347; Vitellinogenin_N_1.
DR SMART; SM00638; LPD_N_1.
DR SMART; SM00216; VWD_1.
KW Glycoprotein; Phosphorylation; Storage protein; signal.
FT SIGNAL 1 15
FT CHAIN 16 1912
FT CHAIN 16 1139
FT CHAIN 1140 1391
FT CHAIN 1392 1627
FT CHAIN 1628 1912
FT DOMAIN 1093 1104
FT DOMAIN 1150 1299
FT DOMAIN 1359 1397
FT DOMAIN 1612 1615
FT CARBOHYD 163 163
FT CARBOHYD 991 991
FT CARBOHYD 1206 1206
FT CARBOHYD 1375 1375
FT CARBOHYD 1662 1662
FT CARBOHYD 1698 1698
FT CARBOHYD 1703 1703
SQ SEQUENCE 1912 AA; 210630 MW; 89BA6273D6492E51 CRC64;

Query Match 4.0%; Score 108; DB 1; Length 1912;
Best Local Similarity 20.3%; Pred. No. 15;
Matches 73; Conservative 39; Mismatches 137; Indels 110; Gaps 10;

OY 176 EQVKVLFEDDERSTSSQCSSEDEIFETPAQVSPRGKQKQARMPLSVQPVSN 235
DB 1076 KRVKKILDDTNGQATRNSSSSASSISSESESTSTSPSSSDSDNRAS----- 1124

```



```

CC      Name=Mei1-1;
CC      IsoId=PF7937-1; Sequence=Displayed;
CC      Name=Mei1-2;
CC      IsoId=PF7937-2; Sequence=VSP 002241;
CC      -1- SIMILARITY: BELONGS TO THE TALE/MEIS HOMEOBOX FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; U68386; AAB19196.1; -
CC      EMBL; U68387; AAB19197.1; -
CC      DR      HSSP; P41778; 1D06.
CC      DR      TRANSFAC; T03390; -
CC      DR      TRANSFAC; T03391; -
CC      DR      InterPro; IPR001356; Homeobox.
CC      DR      Pfam; PF00046; homeobox; 1.
CC      DR      ProDom; PD000010; Homeobox; 1.
CC      DR      SMART; SMO0389; HOX; 1.
CC      DR      PROSITE; PS00027; HOMEOBOX_1; FALSE_NEG.
CC      DR      PROSITE; PS50071; HOMEOBOX_2; 1.
CC      KW      DNA-binding; Nuclear protein; Homeobox; Alternative splicing.
CC      FT      DOMAIN 194 240
CC      FT      DOMAIN 242 269
CC      FT      DNA BIND 272 334
CC      FT      VARSPLIC 373 390
CC      FT      FT
CC      FT      FT
CC      FT      VARIANT 34 34
CC      FT      VARIANT 91 91
CC      FT      VARIANT 178 178
CC      FT      VARIANT 273 273
CC      FT      VARIANT 275 275
CC      FT      VARIANT 315 315
CC      FT      SEQUENCE 390 AA; 42887 MW; DDD98637D08CF32F CRC64;
CC      Query Match 3.9%; Score 106; DB 1; Length 390;
CC      Best Local Similarity 19.2%; Pred. No. 2.4;
CC      Matches 56; Conservative 51; Mismatches 94; Indels 90; Gaps 12;
CC      Oy      181 VFEDDERSTDSQQ-----CSSE--DEDIFETAOVSPRGKEREQWRARMLLSV 230
CC      Db      89 LIFEKCELATCTPRRPVAGVCSSESFNEDIANFSQI-----RAEKPLSS 137
CC      Oy      211 QPVSNADWVWV-----KRLKLCMELCNNTY-----QMLDL-----ENCME 268
CC      Db      138 NPELNLNLIOAIVLRFHLELEKVEHLCDFNCHRYICLEGKMPIDVIDDRDGSKSD 197
CC      Oy      269 EPIFGKGPFFILPFGQSE-----SSTPTGSGSGKETSEDDRSGSRHMEES 317
CC      Db      198 SEDLFRAPLTDQPSWRDHDADASIRSGGTGPGSSGHT--SHSGDNSSSQGGGLNS 254
CC      Oy      318 LSLKAGGDLPLPSPKVEKDSRRKKEWENAKKIYTMADKTIISKLMTEYKKRKQOH 377
CC      Db      255 IASPTGSD-----DDPDKEKK--RKKGIGIFPKXATNIMRAMLF-----QH 294
CC      Oy      378 NLSAFPKVEKVEKKEPKIG-----PRGQSDPLQRPQHLMDQG 415
CC      Db      295 LTHPYPSBEOKKQLAQDTGLTLIQVNNWFNARRIVOPMIDQSNRAVSQG 345
CC      RESULT 12
CC      ANCL1 CAEEL STANDARD; PRT; 8545 AA.
CC      AC Q9N4M4; O61842;
CC      DT 15-SEP-2003 (Rel. 42, Created)

```

```

DT      15-SEP-2003 (Rel. 42, Last sequence update)
DT      15-SEP-2003 (Rel. 42, Last annotation update)
DE      Nuclear anchorage protein 1 (Anchorage 1 protein) (Nesprin homolog).
GN      ANC-1 OR ZK973.6 OR T03A1.3/T03A1.4.
OS      Caenorhabditis elegans.
OC      Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC      Rhabditidae; Peloderiinae; Caenorhabditis.
OX      NCBI_TaxID=6239;
RN      [1]
RP      SEQUENCE OF 1-1751 FROM N.A., FUNCTION, CHARACTERIZATION, AND
RP      INTERACTION WITH F-ACTIN AND UNC-84.
RC      STRAIN=Bristol N2;
RX      MEDLINE=22264052; PubMed=12169658;
RA      Start D.A., Han M.;
RT      "Role of ANC-1 in tethering nuclei to the actin cytoskeleton.";
RT      Science 298:406-409(2002).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=Bristol N2;
RA      Gattung S., Goela D., Broly M.;
RL      Submitted (May-1999) to the EMBL/GenBank/DBJ databases.
RN      [3]
RP      REVISIONS.
RA      Waterston R.;
RL      Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN      [4]
RP      SEQUENCE OF 689-1243 FROM N.A.
RX      MEDLINE=222696983; PubMed=12408964;
RA      Zhang Q., Ragnauth C., Greener M.J., Shanahan C.M., Roberts R.G.;
RT      "The nesprins are giant actin-binding proteins, orthologous to
RT      Drosophila melanogaster muscle protein MSP-300.";
RL      Genomics 80:473-481(2002).
CC      -1- FUNCTION: Plays a central role in nuclear and mitochondrial
CC      anchoring. Probably connects nuclei to the cytoskeleton by
CC      interacting with unc-84 at the nuclear envelope and with F-actin
CC      in the cytoplasm, creating a bridge across the nuclear envelope
CC      between the cytoskeleton and the nucleus.
CC      -1- SUBUNIT: Interacts with F-actin via its N-terminal domain.
CC      -1- SUBCELLULAR LOCATION: Type IV membrane protein (potential). The
CC      largest part of the protein is cytoplasmic, while its C-terminal
CC      part is associated either with the nuclear envelope, most probably
CC      the outer nuclear membrane, or with mitochondrial membrane.
CC      -1- TISSUE SPECIFICITY: Ubiquitously expressed in all postembryonic
CC      cells.
CC      -1- DEVELOPMENTAL STAGE: First expressed in L1 larvae, and thereafter.
CC      -1- DOMAIN: The large coiled coil domains are composed of 6 nearly
CC      exact repeats of 903 residues. The last repeat is partial. These
CC      repeats are conserved in Hawai (CB4856), Australia (AB4) and
CC      Germany (RC301) strains. The length of the repeat may be
CC      maintained because of a selective advantage to keep the protein
CC      large and allow a single molecule to extent more than 0.5
CC      micrometers.
CC      -1- DOMAIN: The Karsicht domain, which contains a potential
CC      transmembrane domain, is essential for the nuclear envelope
CC      targeting.
CC      -1- SIMILARITY: Belongs to the Nesprin family.
CC      -1- SIMILARITY: Contains 1 actin-binding domain.
CC      -1- SIMILARITY: Contains 2 calponin-homology (CH) domains.
CC      -1- SIMILARITY: Contains 1 Karsicht domain.
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; AY157938; AAN35200.1; -
CC      DR      EMBL; AC006834; AAF40010.3; -
CC      DR      EMBL; AY126454; AAM95163.1; -
CC      HSSP; P46939; 10QG.

```

DR WormPep; ZK973.6; CE33588.
 DR InterPro; IPR001589; Actbind_actuin.
 DR InterPro; IPR001715; Calponin-like.
 DR InterPro; IPR002017; Spectrin.
 DR Pfam; PF00307; CH; 2.
 DR SMART; SM00264; BAG; 6.
 DR SMART; SM00033; CH; 2.
 DR PROSITE; PS00019; ACTININ_1; 1.
 DR PROSITE; PS00020; ACTININ_2; FALSE_NEG.
 DR PROSITE; PS50021; CH; 2.
 KM Structural protein; Cytoskeleton; Actin-binding; Transmembrane;
 KM Coiled coil; Repeat.
 FT TRANSMEM 8495 8513
 FT DOMAIN 1 8494
 FT DOMAIN 8514 8545
 FT DOMAIN 1 325
 FT DOMAIN 23 130
 FT DOMAIN 222 325
 FT DOMAIN 3241 8199
 FT REPEAT 3241 4143
 FT REPEAT 4144 5097
 FT REPEAT 5098 6000
 FT REPEAT 6001 6803
 FT REPEAT 6904 7806
 FT REPEAT 7807 8199
 FT DOMAIN 754 774
 FT DOMAIN 1072 1101
 FT DOMAIN 1215 1236
 FT DOMAIN 1124 1184
 FT DOMAIN 1574 1629
 FT DOMAIN 1725 1754
 FT DOMAIN 1950 1981
 FT DOMAIN 2103 2580
 FT DOMAIN 2582 2712
 FT DOMAIN 2852 2949
 FT DOMAIN 3002 3119
 FT DOMAIN 3178 3295
 FT DOMAIN 3346 3417
 FT DOMAIN 3482 3552
 FT DOMAIN 3587 3703
 FT DOMAIN 3781 3839
 FT DOMAIN 3902 4022
 FT DOMAIN 4114 4198
 FT DOMAIN 4249 4320
 FT DOMAIN 4436 4506
 FT DOMAIN 4541 4657
 FT DOMAIN 4735 4793
 FT DOMAIN 4856 4976
 FT DOMAIN 5035 5152
 FT DOMAIN 5203 5274
 FT DOMAIN 5339 5409
 FT DOMAIN 5444 5560
 FT DOMAIN 5638 5696
 FT DOMAIN 5759 5879
 FT DOMAIN 5938 6055
 FT DOMAIN 6106 6177
 FT DOMAIN 6242 6312
 FT DOMAIN 6347 6463
 FT DOMAIN 6541 6599
 FT DOMAIN 6662 6782
 FT DOMAIN 6841 6958
 FT DOMAIN 7009 7080
 FT DOMAIN 7145 7215
 FT DOMAIN 7250 7366
 FT DOMAIN 7444 7502
 FT DOMAIN 7565 7685
 FT DOMAIN 7744 7861
 FT DOMAIN 7912 7983
 FT DOMAIN 8048 8118
 FT DOMAIN 8153 8204
 FT DOMAIN 8273 8329
 FT DOMAIN 8370 8390
 COILED COIL (POTENTIAL).
 COILED COIL (POTENTIAL).
 ANCHOR FOR TYPE IV MEMBRANE PROTEIN
 PERINUCLEAR SPACE (POTENTIAL).
 CH 1.
 CH 2.
 6 X TANDEM REPEAT.
 1.
 2.
 3.
 4.
 5.
 6.

FT DOMAIN 8486 8545 KLARSICHT.
 FT DOMAIN 1655 1658 POLY-ARG.
 FT DOMAIN 3021 3029 POLY-LYS.
 FT DOMAIN 3924 3932 POLY-LYS.
 FT DOMAIN 4878 4886 POLY-LYS.
 FT DOMAIN 5781 5789 POLY-LYS.
 FT DOMAIN 6684 6692 POLY-LYS.
 FT DOMAIN 7587 7595 POLY-LYS.
 SO SEQUENCE 8545 AA; 956470 MM; 02A94D994BEE19E3 CRC64;
 Query Match 3.9%; Score 106; DB 1; Length 8545;
 Best Local Similarity 18.6%; Pred. No. 1.se+02;
 Matches 100; Conservative 83; Mismatches 196; Indels 158; Gaps 23;
 QY 14 TQSPKTPNNPDHAGSQQLIIELPDEKPRGHT---KKSVSPREIVSLSHQVLLDNLY 70
 DB 3582 TQVAPKAEELKEVEAKAVIANNPS--DAHVQQLQAVATVETLLIPDEERARLNNF 3639
 QY 71 -----DILLEEF-----VKGPSPEEKTIOVPE-----AKLAGFLRYI 103
 DB 3640 LAARNDIDALIBQLQPLDAVLAQPKRSABEAQDVENLANNSQLSDLNKIA----- 3693
 QY 104 SMQNLAVIFDLLDSTYRTAEPFTSPGLKCLKKVSIGGAAULYROSAMFNIFYHALV 163
 DB 3694 NLQRISELDPLEBAVDVRFPPD----- 3716
 QY 164 CAVLTNOETITAEQVK---KVL-----FEDDERSTDSSQCSSEDEDIPEETAQVSP 213
 DB 3717 -----VDAETTRQYDVLNDVAELDEETLLKQSAQVANNETIDISKMIDSTDPE 3767
 QY 214 RGKEKRWARMPLLSVO-----PVSNAQVWMLVKRLH-----KLCMEICNNYIQMHL 262
 DB 3768 RSLIDTAKSDIPALKAQINRIKRIYVAD---ASKKHVTTOPKIAEDLDNKLAKLQTE 3823
 QY 263 LENCM-----EPPPI--FKGDP-----PLPSPQSESTPST 293
 DB 3824 LDDAIKTSDEHDKQQLLSLKNISQFEQIPLDQLKSDDKTAKETNSLKEPEAEPL 3883
 QY 294 GGFSG-KEPSPEDRSQREHMGESLSLKAGGDLPLPSPKYEKQPSRK--KEWMENAG 351
 DB 3884 AKIQELREAKRVGGEAASAAH-DQIVALEKEADLVTKESAKKKKKKKSPQEMTELS 3942
 QY 352 NKITYMAADTKISK-----LMTEYKRRKQO--HNLAPPKXV--KYEKKEPPLGPR 398
 DB 3943 AKV--VEAKKLLPIEEBAKKNENLPADKPKAEGLVNLAFYQVDVETQSEKQDELDKL 4000
 QY 399 GQDSPLLQRPQHLMDQCMHSPSAGPELLRQDKRPSPSGTSGSLSVVPDAEAQIQ 455
 DB 4001 NNANDAIKRLGDLDDAEKTVVPSVPAL--SEFKRIAPHLATLVAVNDVPASVE 4055
 RESULT 13
 ID GTA_NPVAC STANDARD; PRT; 506 AA.
 AC P41447;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DS Probable global transactivator.
 GN GTA.
 OS Autographa californica nuclear polyhedrosis virus (AcMNPV).
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolyhedroviruses.
 OX NCBI_Taxid=46015;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C6;
 RX MEDLINE=94303173; PubMed=8030224;
 RA Ayres M.D., Howard S.C., Kuzio J., Lopez-Farber M., Possee R.D.;
 RT "The complete DNA sequence of Autographa californica nuclear
 polyhedrosis virus."
 RL Virology 202:586-605(1994).
 CC -!- SIMILARITY: BELONGS TO THE SNF2/RAD54 HELICASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
CC EMBL: L22858; AAA66672.1; -
CC PIR: B72855; B72855.
CC InterPro: IPR001410; DEAD.
CC InterPro: IPR001650; Helicase_C.
CC InterPro: IPR000330; SNF2_N.
CC Pfam: PF00271; helicase_C; 1.
CC SMART: PF00176; SNF2_N; 1.
CC SMART: SM00487; DEXDC; 1.
CC SMART: SM00490; HELICC; 1.
CC Helicase; ATP-binding.
CC NP_BIND 58 64 ATP (BY SIMILARITY).
CC SITE 157 160 DEAD BOX.
CC SEQUENCE 506 AA; 59058 MW; 8480CC24A85984AE CRC64;

Query Match 3.9%; Score 105; DB 1; Length 506;
Best Local Similarity 22.3%; Pred. No. 4.1;
Matches 69; Conservative 49; Mismatches 91; Indels 100; Gaps 18;

QY 126 DTPSLKLLKKVSGIGGAAALYR-QSAMSFNIFHALVCAVLTNQETITAEQVKVFE 184
DB 222 DSTNRKIKIKIV-----LKRKSEISNIPGH-----TYEYV-HVNFN 260
QY 185 DDERSTDSQCCSDEDEIFETAYQVSPRGKEKQKQKARMPILSVQVSNADWVLYKR 244
DB 261 EEEKLYDKLKESE-----AYKAVAAARENENALSR-LQOMOHV-----LMLIK 306
QY 245 LHKLMELCNNTYIOMHLDLENCEPPIFKGDPF-----FILPS 283
DB 307 LKQIC--CHPYLANH--GKNILETNDCTKMDYMSKCKRYLDVLDILNLSNDKILLVS 361
QY 284 -----FQS--ESSTPSYTGFSK-----ETPSDDRSGSRHMGESLSLKAG 324
DB 362 QWVEYLIKIFENFKQNIATLMTYGLKVEDRILETTFND--AANTQHRILLISIKGG 419
QY 335 GDL-----LLEP--SPKVEKQDPSRKKEWENAGKITYM--AADKITSKMTBYK 371
DB 420 VGLNLIGSNHLYMELPHWNPQLEIQAQDRISMGGTKQTYVYKMLNVEDNSIEK----YI 475
QY 372 KRKQOHNS 380
DB 476 KORQDKIA 484

RESULT 14
BRF3 HUMAN
ID BRF3 HUMAN STANDARD; PRT: 1214 AA.
AC O9UUD4:
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Bromodomain and PHD finger-containing protein 3 (Fragment).
GN BRPF3 OR KIAA1286.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20039619; PubMed=10574462;
RA Nagase T., Ishikawa K.-I., Kikuno R., Hirose M., Nomura N.,
RA Ohara O.,
RT "Prediction of the coding sequences of unidentified human genes. XV.
RT The complete sequences of 100 new cDNA clones from brain which code

RT for large proteins in vitro.";
RL DNA Res. 6:337-345 (1999).
CC -1- SIMILARITY: Contains 1 bromodomain.
CC -1- SIMILARITY: Contains 1 PHD-type zinc finger.
CC -1- SIMILARITY: Contains 1 PHD-type zinc finger.
CC -1- SIMILARITY: Contains 1 PHD-type zinc finger.
CC -1- SIMILARITY: Contains 1 PHD-type zinc finger.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
CC EMBL: AB033112; BAA86600.1; -
CC HSSP: Q92831; 1B91.
CC Genew: HGNC:14256; BRPF3.
CC InterPro: IPR001487; Bromodomain.
CC InterPro: IPR000313; PWWP domain.
CC InterPro: IPR001965; Znf_Phd.
CC Pfam: PF00439; bromodomain; 1.
CC Pfam: PF00628; PHD; 1.
CC Pfam: PF00855; PWWP; 1.
CC PRINTS: PR00503; BROMODOMAIN.
CC SMART: SM00297; BROMO; 1.
CC SMART: SM00249; PHD; 2.
CC SMART: SM00293; PWWP; 1.
CC PROSITE: PS00633; BROMODOMAIN_1; FALSE_NEG.
CC PROSITE: PS50014; BROMODOMAIN_2; 1.
CC PROSITE: PS01359; ZF_Phd_1; 1.
CC PROSITE: PS50016; ZF_Phd_2; 1.
CC Zinc-finger; Bromodomain.
CC NON_TER 1
CC ZN_FING 221 271 PHD-TYPE.
CC DOMAIN 415 441 GLU-RICH.
CC DOMAIN 615 685 BROMODOMAIN.
CC DOMAIN 1085 1168 PWWP.
CC SEQUENCE 1214 AA; 136598 MW; CA490810622109CD CRC64;

Query Match 3.9%; Score 105; DB 1; Length 1214;
Best Local Similarity 22.5%; Pred. No. 13;
Matches 109; Conservative 65; Mismatches 181; Indels 130; Gaps 27;

QY 28 QSCQILIELPDEKPNKGTGKVSFREIVSLSHQVLYLLEEFVGGSPGEK 87
DB 569 ERARLILIELRKREKLRQGVQVQAAAMELELMPFNILRTTLDLQ--KDAHIFAE 625
QY 88 TIQVPEAKLAGFLAYIGMQLAVIFDLDS--YRTAREPTSGKCLKKVSGIGAA 145
DB 626 PNLSE--VPDYLEFISKPMDFSTMRKLSHLRYLLEFEED-----F 667
QY 146 NLVRSAMSFN-----IFYHALV-----CAVLTNQETITAEQVKVLEFEDERSTDSQ 195
DB 668 NLIYTNCKYKVAKOTIFRAAVRLDGLGAILRH--ARQENIYDPERGTHLPES 722
QY 196 CSSED-----ED-----IFETAYQVSPRGKEKQKQKARMPILSVQVSNADWVLYKR 244
DB 723 PKLDEPFYRSWEDVDNLIIPENRALSP--VQLKELKLDLSAMRSSGA--RTRR 776
QY 245 LHKLMELCNNTYIOMHLDLENCEPPIFKGDPPIFLPSPOSESTSTG--GFGS-- 298
DB 777 VRLIRREI--NALQOKL--ADPP-----PQPSLTKTVSNGSLPAGQGDA 819
QY 299 -----KETPSD--DRSQSR-----EHMGESLSLKAGGDLILPPSPV-----EKK 338
DB 820 AVLQALQEBEDDGDGDDSLPPPLLEPRGAPSL-----SEGESPEPPTLKPIINDSK 875
QY 339 DPSR-----KKEWENA--GNK-IYTMADKITSKLMTEYKRRKQOHNSAPFKE 385
DB 876 PPSRFLKPRKVEDELEKSPQLQGNBPLOQLISDNGINRLSLMAPTPAGTPTGCGRR 935
QY 386 VKVEKKEPRLGPRQDPSLQD--PQHLMQDQGMHNSAGP-----ELLRQDKRRPSG 437

```

Db      936 TSV-----LTKKAKNGVKLQSRSPDVLNCE-DHGVAGSPASPASTEERHSKRPRSR 988
Qy      438 STGS 442
Db      989 SCSES 993

RESULT 15
VP41_CAEEL STANDARD; PRT; 876 AA.
ID      019954
AC      01-NOV-1997 (Rel. 35, Created)
DT      28-FEB-2003 (Rel. 41, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Vacuolar assembly protein VPS41 homolog.
GN      F32A6.3.
OS      Caenorhabditis elegans.
OC      Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC      Rhabditidae; Peloderinae; Caenorhabditis.
OX      NCBI_Taxid=6239;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=Bristol NZ;
RA      Wu X.;
RL      Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
RP      REVISIONS.
RA      Waterston R.;
RL      Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC      -1- FUNCTION: REQUIRED FOR VACUOLAR ASSEMBLY AND VACUOLAR TRAFFIC (BY
CC      SIMILARITY).
CC      -1- SIMILARITY: BELONGS TO THE VPS41 FAMILY.
CC      -1- SIMILARITY: Contains 1 RING-type zinc finger.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; U40409; AAA81388.2; -.
DR      Wormpep; F32A6.3; CE29305.
DR      InterPro; IPR000547; Clathrin_repeat.
DR      InterPro; IPR001841; Znf_ring.
DR      Pfam; PF00637; Clathrin; 1.
DR      SMART; SM00239; CLH; 1.
DR      SMART; SM00184; RING; 1.
DR      PROSITE; PS00518; ZF_RING_1; FALSE_NEG.
DR      PROSITE; PS50089; ZF_RING_2; 1.
KM      Hypothetical protein; Zinc-finger.
FT      ZN_FING 814 868 RING-TYPE.
SQ      SEQUENCE 876 AA; 99987 MM; 8D4B0E2AD4211C06 CRC64;

Query Match 3.9%; Score 104.5; DB 1; Length 876;
Best Local Similarity 18.3%; Pred. No. 9.2; Indels 93; Gaps 15;
Matches 66; Conservative 64; Mismatches 138;

Qy      1 RIRAAQOVFMIDQCSPK---TPNPFDAQSCQILIELPPDEKPNHGHTKKSVSFREIVV 57
Db      425 RARAASRLPLICGCKREWEVAVQFEVGLCTLLAEVLDPDGP---TLDECYQKVL 481
Qy      58 SLASHOV-----LIQONLY-DILLEEFV-----KQSPBGEKTIQVPAKLAG 98
Db      482 ACLFNINVKQFRKLVQTMSPDLYMTSFFIDRTQWRIOQISKSGNLADVDETERVLMADALAH 541
Qy      99 FLRY-----ISMONTAV-----IFDLILDSYRTAREFDTS-PGLKCLKKVS 139
Db      542 LYLIERKESALKLIMSQDQFOIFNVLDKQGLFDLVKQDITELMNINSEKALRLLDNAD 601
Qy      140 GIGGA---ANLYROSAHSFNIFYFHALVCAVL-----NQETITAEQVKKVLFEDDERS-- 189

```

```

Db      602 SVSPSFVMEKIGROPKQLQ-----AYLTKLMSRNEGTETFPADKAVQLYAEYDQCKLL 652
Qy      190 -----TSSQCCSEDEDIFETTAQVSPPRGKE-----KROMRAMPILISVQ 231
Db      653 PFLRKNNANVNVKARKLCS--DKGYIEETIYLAKSGNHYDAVKMNVREYRNNEKVIDYC 710
Qy      232 PVSNAQVWVLVKRLKLCWELCNVYIQMHLDLNCMEEPPIFKGDPPFLLPSFQSSSTP 291
Db      711 KDQNDPDLWT--HILGVVAEFPAPHFSQLIEASNCL-----DPLINDKLPDSDIDP 760
Qy      292 S 292
Db      761 N 761

```

Search completed: July 24, 2003, 11:54:17
Job time : 28.511 secs

THIS PAGE BLANK (u970)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 11:47:53 ; Search time 115.28 Seconds
(without alignments)
1159.539 Million cell updates/sec

Title: US-09-991-681-27
Perfect score: 2698
Sequence: 1 RIRMAQOVFMIDTQCSPKT.....VROAVREWLGRVGRVYDIIV 518

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2698	100.0	1770	4	Q9ULH6	Q9ULH6 homo sapien
2	2675.5	99.2	589	4	Q8N4Y4	Q8N4Y4 homo sapien
3	2675.5	99.2	592	4	Q96CH9	Q96CH9 homo sapien
4	177.5	6.6	2016	5	Q9W444	Q9W444 drosophila
5	133	4.9	687	5	Q9VXT5	Q9VXT5 drosophila
6	130.5	4.8	751	5	Q9SR63	Q9SR63 drosophila
7	129.5	4.8	1330	5	Q9VJ87	Q9VJ87 drosophila
8	124	4.6	2150	5	Q17596	Q17596 caenorhabdi
9	122	4.5	809	5	Q961X4	Q961X4 drosophila
10	122	4.5	1657	5	Q9VU76	Q9VU76 drosophila
11	116.5	4.3	1433	3	Q8TF17	Q8TF17 saccharomyc
12	116	4.3	1110	11	Q8CDI7	Q8CDI7 mus musculu
13	115	4.3	2309	3	Q8WZV3	Q8WZV3 neurospora
14	114	4.2	1077	4	Q8TE71	Q8TE71 homo sapien
15	112	4.2	565	13	Q8J1P0	Q8J1P0 oryzae lat
16	111	4.1	594	4	Q8TAR2	Q8TAR2 homo sapien

17	111	4.1	594	4 Q92529	Q92529 homo sapien
18	111	4.1	781	3 Q74526	Q74526 echinosacch
19	111	4.1	1148	4 Q99590	Q99590 homo sapien
20	110	4.1	794	5 Q9W566	Q9W566 drosophila
21	110	4.1	794	5 Q9SR31	Q9SR31 drosophila
22	109.5	4.1	587	5 Q21893	Q21893 caenorhabdi
23	108.5	4.0	886	4 Q8N2Y6	Q8N2Y6 homo sapien
24	108.5	4.0	897	4 Q9BWD2	Q9BWD2 homo sapien
25	108.5	4.0	962	5 Q9VTE2	Q9VTE2 drosophila
26	108	4.0	506	12 Q92409	Q92409 bombyx mori
27	108	4.0	1105	5 Q94314	Q94314 caenorhabdi
28	108	4.0	1583	3 Q8N1X2	Q8N1X2 neurospora
29	107.5	4.0	2282	4 Q8NFW7	Q8NFW7 homo sapien
30	107.5	4.0	2284	5 Q8TIV3	Q8TIV3 dicystoselei
31	107.5	4.0	2385	4 Q8IVL0	Q8IVL0 homo sapien
32	107	4.0	537	11 Q91XR3	Q91XR3 mus musculu
33	107	4.0	537	11 Q8BHL3	Q8BHL3 mus musculu
34	107	4.0	714	11 Q8K2L9	Q8K2L9 mus musculu
35	107	4.0	829	6 Q95KJ9	Q95KJ9 macaca fasc
36	107	4.0	983	4 Q9COA4	Q9COA4 homo sapien
37	107	4.0	1337	4 Q9Y2L2	Q9Y2L2 homo sapien
38	107	4.0	2724	4 Q96SE4	Q96SE4 homo sapien
39	107	4.0	3467	5 Q81218	Q81218 plasmodium
40	106.5	3.9	1311	4 Q81ZX7	Q81ZX7 homo sapien
41	106.5	3.9	1988	5 Q9VU16	Q9VU16 drosophila
42	106.5	3.9	2617	16 Q8YTR9	Q8YTR9 anabaena sp
43	106	3.9	522	2 Q937U4	Q937U4 bacteroides
44	106	3.9	522	2 Q8G9E2	Q8G9E2 bacteroides
45	106	3.9	661	10 Q94CY0	Q94CY0 oryza sativ

ALIGNMENTS

RESULT 1

Q9ULH6 PRELIMINARY; PRT; 1770 AA.

AC Q9ULH6; Q96P46; RT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)

DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)

DE BIG3.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Hong W.;

RT "KIAA1244 as a novel distantly related member (BIG3) of the BIG1/Sec7p

RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF413080; ALA04174.1; -.

DR InterPro; IPR000904; Sec7.

DR SMART; SM00222; Sec7; 1.

KW Hypothetical protein.

SQ SEQUENCE 1770 AA; 195845 MW; 5E996E36A6F92ABA CRC64;

Query Match 100.0%; Score 2698; DB 4; Length 1770;
Best Local Similarity 100.0%; Pred. No. 1.4e-225;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	RIRMAQOVFMIDTQCSPKTNNPDHNSCOLIIEPPDEKPGHTKSSVFEIVSL 60
DB	1253	RIRMAQOVFMIDTQCSPKTNNPDHNSCOLIIEPPDEKPGHTKSSVFEIVSL 1312
QY	61	SHOVLQNTLYDILIEEFVKGSPGSEKTIQVPEAKLAGFLRYISMONTAVIPDLLDSYR 120
DB	1313	SHOVLQNTLYDILIEEFVKGSPGSEKTIQVPEAKLAGFLRYISMONTAVIPDLLDSYR 1372
QY	121	TAREPDTSPGLKLLKLVSGIGGANLYROSAMSFNIYFHALVCAYVTNDETTTAEQVK 180
DB	1373	TAREPDTSPGLKLLKLVSGIGGANLYROSAMSFNIYFHALVCAYVTNDETTTAEQVK 1432

```

OY 181 VLFEDERSSTDSQCCSEDEDIFEEITAOVSPPRKKEKQMRAMPPLLSVQPVSNADWW 240
DB 1433 VLFEDERSSTDSQCCSEDEDIFEEITAOVSPPRKKEKQMRAMPPLLSVQPVSNADWW 1492
OY 241 LVKRLHKLCEMLCNVYIQMHLDLENCEMBEPIFKGDPPFILPSPQSSSTPTSGFSGKE 300
DB 1493 LVKRLHKLCEMLCNVYIQMHLDLENCEMBEPIFKGDPPFILPSPQSSSTPTSGFSGKE 1552
OY 301 TPSEDDRSQSRHEHGESISLKAGGDDLPLPSPKYEKQDPSRKKEWENAGNKITYMAAD 360
DB 1553 TPSEDDRSQSRHEHGESISLKAGGDDLPLPSPKYEKQDPSRKKEWENAGNKITYMAAD 1612
OY 361 KTISKLMEYKKRKQOHNLSAFPKEVKEKKEBPGRGQDSPILQRPQHLMDGQMRHS 420
DB 1613 KTISKLMEYKKRKQOHNLSAFPKEVKEKKEBPGRGQDSPILQRPQHLMDGQMRHS 1672
OY 421 FSAGPELLRODKRPRSGSTGSSLSVSVDAEAQIQAMTNMVLTVLNOIQILPDQFTALQ 480
DB 1673 FSAGPELLRODKRPRSGSTGSSLSVSVDAEAQIQAMTNMVLTVLNOIQILPDQFTALQ 1732
OY 481 PAVFPCISQLTCHVTDIRVROAVREMLGRVGRVYDIIV 518
DB 1733 PAVFPCISQLTCHVTDIRVROAVREMLGRVGRVYDIIV 1770

RESULT 2
OEN4Y4 PRELIMINARY; PRT; 589 AA.
AC OEN4Y4;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DB Similar to KIAA1244 protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
CX NCBI_TaxId=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC033191; AAH3191.1; -.
FT NON TER
SQ SEQUENCE 589 AA; 66086 MW; 8041EEA348DE65F7 CRC64;

Query Match 99.2%; Score 2675.5; DB 4; Length 589;
Best Local Similarity 99.6%; Pred. No. 2.6e-224;
Matches 516; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
OY 1 RIRAMAQOVFMDLTQCSPTKPNPFDAOSCOLIIELPDEKPNGHTKKSVSFREIVSIL 60
DB 73 RIRAMAQOVFMDLTQCSPTKPNPFDAOSCOLIIELPDEKPNGHTKKSVSFREIVSIL 132
OY 61 SHQVLLQNLVYDILLEEFVKGPSPEBEKTIQVPEAKLAGFLRYISQNLAVIFDILLDSYR 120
DB 133 SHQVLLQNLVYDILLEEFVKGPSPEBEKTIQVPEAKLAGFLRYISQNLAVIFDILLDSYR 192
OY 121 TAREFDTPSGKLCILKKYSGIGGANLYROSAMSFNIYFHALVCAVLTNOETTTAEQYKK 180
DB 193 TAREFDTPSGKLCILKKYSGIGGANLYROSAMSFNIYFHALVCAVLTNOETTTAEQYKK 252
OY 181 VLFEDERSSTDSQCCSEDEDIFEEITAOVSPPRKKEKQMRAMPPLLSVQPVSNADWW 240
DB 253 VLFEDERSSTDSQCCSEDEDIFEEITAOVSPPRKKEKQMRAMPPLLSVQPVSNADWW 311
OY 241 LVKRLHKLCEMLCNVYIQMHLDLENCEMBEPIFKGDPPFILPSPQSSSTPTSGFSGKE 300
DB 312 LVKRLHKLCEMLCNVYIQMHLDLENCEMBEPIFKGDPPFILPSPQSSSTPTSGFSGKE 371
OY 301 TPSEDDRSQSRHEHGESISLKAGGDDLPLPSPKYEKQDPSRKKEWENAGNKITYMAAD 360

```

```

DB 372 TPSEDDRSQSRHEHGESISLKAGGDDLPLPSPKYEKQDPSRKKEWENAGNKITYMAAD 431
OY 361 KTISKLMEYKKRKQOHNLSAFPKEVKEKKEBPGRGQDSPILQRPQHLMDGQMRHS 420
DB 432 KTISKLMEYKKRKQOHNLSAFPKEVKEKKEBPGRGQDSPILQRPQHLMDGQMRHS 491
OY 421 FSAGPELLRODKRPRSGSTGSSLSVSVDAEAQIQAMTNMVLTVLNOIQILPDQFTALQ 480
DB 492 FSAGPELLRODKRPRSGSTGSSLSVSVDAEAQIQAMTNMVLTVLNOIQILPDQFTALQ 551
OY 481 PAVFPCISQLTCHVTDIRVROAVREMLGRVGRVYDIIV 518
DB 552 PAVFPCISQLTCHVTDIRVROAVREMLGRVGRVYDIIV 589

RESULT 3
O96CH9 PRELIMINARY; PRT; 592 AA.
AC O96CH9;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DB Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
CX NCBI_TaxId=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC014227; AAH14227.1; -.
KW Hypothetical protein.
FT NON TER
SQ SEQUENCE 592 AA; 66400 MW; FA1E807B0DF47B5 CRC64;

Query Match 99.2%; Score 2675.5; DB 4; Length 592;
Best Local Similarity 99.6%; Pred. No. 2.6e-224;
Matches 516; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
OY 1 RIRAMAQOVFMDLTQCSPTKPNPFDAOSCOLIIELPDEKPNGHTKKSVSFREIVSIL 60
DB 76 RIRAMAQOVFMDLTQCSPTKPNPFDAOSCOLIIELPDEKPNGHTKKSVSFREIVSIL 135
OY 61 SHQVLLQNLVYDILLEEFVKGPSPEBEKTIQVPEAKLAGFLRYISQNLAVIFDILLDSYR 120
DB 136 SHQVLLQNLVYDILLEEFVKGPSPEBEKTIQVPEAKLAGFLRYISQNLAVIFDILLDSYR 195
OY 121 TAREFDTPSGKLCILKKYSGIGGANLYROSAMSFNIYFHALVCAVLTNOETTTAEQYKK 180
DB 196 TAREFDTPSGKLCILKKYSGIGGANLYROSAMSFNIYFHALVCAVLTNOETTTAEQYKK 255
OY 181 VLFEDERSSTDSQCCSEDEDIFEEITAOVSPPRKKEKQMRAMPPLLSVQPVSNADWW 240
DB 256 VLFEDERSSTDSQCCSEDEDIFEEITAOVSPPRKKEKQMRAMPPLLSVQPVSNADWW 314
OY 241 LVKRLHKLCEMLCNVYIQMHLDLENCEMBEPIFKGDPPFILPSPQSSSTPTSGFSGKE 300
DB 315 LVKRLHKLCEMLCNVYIQMHLDLENCEMBEPIFKGDPPFILPSPQSSSTPTSGFSGKE 374
OY 301 TPSEDDRSQSRHEHGESISLKAGGDDLPLPSPKYEKQDPSRKKEWENAGNKITYMAAD 360
DB 375 TPSEDDRSQSRHEHGESISLKAGGDDLPLPSPKYEKQDPSRKKEWENAGNKITYMAAD 434
OY 361 KTISKLMEYKKRKQOHNLSAFPKEVKEKKEBPGRGQDSPILQRPQHLMDGQMRHS 420
DB 435 KTISKLMEYKKRKQOHNLSAFPKEVKEKKEBPGRGQDSPILQRPQHLMDGQMRHS 494
OY 421 FSAGPELLRODKRPRSGSTGSSLSVSVDAEAQIQAMTNMVLTVLNOIQILPDQFTALQ 480
DB 495 FSAGPELLRODKRPRSGSTGSSLSVSVDAEAQIQAMTNMVLTVLNOIQILPDQFTALQ 554

```

Qy	481	PAVFPICISQTCVTDTRNQAVREMGVGRVYDII	518
Db	555	PAVFPICISQTCVTDTRNQAVREMGVGRVYDII	592
RESULT 4			
ID	Q9M444	PRELIMINARY;	PRT; 2016 AA.
AC	Q9M444		
DT	01-MAY-2000 (TrEMBLrel. 13, Created)		
DT	01-OCT-2002 (TrEMBLrel. 23, Last sequence update)		
DT	01-MAR-2003 (TrEMBLrel. 22, Last annotation update)		
DE	CG5937 protein.		
GN	CG5937		
OS	Drosophila melanogaster (Fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;		
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
OC	Ephydroidea; Drosophilidae; Drosophila.		
NCBI	NCBI_TaxId=7227;		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Berkley;		
RX	MEDLINE=20196006; PubMed=10731132;		
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,		
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,		
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,		
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,		
RA	Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,		
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,		
RA	Abvill J.F., Agbayani A., An H.-J., Andrews-Plankoch C., Baldwin D.,		
RA	Balcer R.M., Baas A., Baxendale J., Bayraktaroglu L., Beasley E.M.,		
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,		
RA	Botkova D., Botchan M.R., Bouck J., Brokstein P., Broctier P.,		
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,		
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,		
RA	de Pabloe B., Delcher A., Deng Z., Deynadt A.D., Dew I., Dietz S.M.,		
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,		
RA	Durbin K.J., Evangelista C.C., Ferraz S.C., Ferreira S., Fleischmann W.,		
RA	Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glaeser K.,		
RA	Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,		
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,		
RA	Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,		
RA	Jatali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,		
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,		
RA	Lau X.P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,		
RA	Lisio X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,		
RA	Merkulov G., Mishina N.V., Mobery C., Morris J., Moshrefi A.,		
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,		
RA	Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pachtel J.M.,		
RA	Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,		
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,		
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,		
RA	Spieler E., Spadling A.C., Stapleton M., Strong R., Sun E.,		
RA	Swirebas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,		
RA	Wang Z.-Y., Wassarman D.A., Weinstein K.W., Weissbach J.,		
RA	Williams S.M., Woodage T., Worley R.C., Wu D., Yang S., Yao Q.A.,		
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,		
RA	Zheng X.H., Zhong F., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,		
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,"		
RT	"The genome sequence of Drosophila melanogaster.";		
RL	Science 287;2185-2195(2000).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RA	Celniker S.E., Adams M.D., Krommler B., Wan K.H., Holt R.A.,		
RA	Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,		
RA	Barzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,		
RA	Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,		
RA	Dodson K., Dorett V., Doup L.E., Doyle C., Drensek D., Farfan D.,		
RA	Ferreira S., Frise E., Galle R.F., Gary N.S., George R.A.,		
RA	Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,		
RA	Ibegwam C., Jatali M., Kruse D., Li P., Mattei B., Moshrefi A.,		
RA	McIntosh T.C., Moy M., Murphy B., Nelson C.A., Nunoo J.,		
RA	Pachtel J., Paragás V., Park S., Patel S., Pfeiffer B.,		

RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Stykars J., Tecor C., Tyler D.,
RA Williams S.M., Zaveri U.S., Smith H.O., Venter J.C., Rubin G.M.,
RT Sequencing of *Drosophila melanogaster* genome. ";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RA Miera S., Crosby M.A., Matthews B.B., Bayraktarolu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Bernan B., Carlson J.W., Celniker S.E.,
RA Clamp M., Dyrasale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Krommiller B., Marshall B., Milburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutnick F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.,
RT "Annotation of *Drosophila melanogaster* genome. ";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[4]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.U.,
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[5]
RP SEQUENCE FROM N.A.
RA Flybase:
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003436; AA046115.2; -
DR FlyBase; FBgn0029834; CG5937.
DR InterPro; IPR002173; PFkB.
DR InterPro; IPR000904; Sec7.
DR SMART; SM00222; Sec7; 1.
DR SMART; PS00583; PFkB_KINASES_1; 1.
DR SEQUENCE; 2016 AA; 223895 MW; 94ED4C6DA7EBB5D0 CRC64;
SQ

Query Match	6.6%	Score 177.5	DB: 5	Length 2016
Best Local Similarity	29.9%	Pred. No. 7,2e-06		
Matches 60	Conservative 25	Mismatches 53	Indels 63	Gaps 7

Qy	1	RIRAAQOVFLMDTQCS	-----KTPNNFDHAQSCOLIILPEPDEK	-----NGAT	46
		:	:	:	
Db	1610	RIVAAQOVFLSDQRPQGNQAPT	PS-----ASQCKL	-----SDDRSYSLFLYPLANN	1660
Qy	47	KK-----SVSFREIVVSLSHQVLLQNLVLDLLEF	-----VKGPS	82	
Db	1661	SNIDNFVIRIPFKXVLVGLLANQMLQVLAKLLSLRKVCVPAQAVSTCI	FDNYAASAPS	1720	
Qy	83	PGEKTIQVBEAKLAGEFLRYISMONTLAVIFDILLDSYRTAREPDTSPGLKCLKLVSGIG	142		
Db	1721	HDYDLDRSRKEILLRCQKQYL	-----MSALFEDSRGLKFLMKVSNIE	1764	
Qy	143	GAANLYROSAMSFNIYFHALV	163		
		:	:		
Db	1765	YAANLYRKQMTSSWMIYIALV	1785		

RESULT 5	Q9VXT5	PRELIMINARY;	PRT;	687	AA.
AC	Q9VXT5				
DT	01-MAY-2000	(Tremblere)	13,	Created	
DT	01-OCT-2002	(Tremblere)	22,	Last sequence update	
DT	01-MAR-2003	(Tremblere)	23,	Last annotation update	
DE	CG32585	protein (LP28117p).			
GN	CG9213.				
OS	Drosophila melanogaster	(fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;				
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;				
OC	Ephydroidea; Drosophilidae; Drosophila.				
OX	NCBI_Taxid=7227;				
OX	NCBI_Taxid=7227;				
RN	SEQUENCE FROM N.A.				
RP	STRAIN=Berkeley;				
RC	MELLINE=20196006;	Pubmed=10731132;			
RA	Adams M.D., Celliker S.E., Holt R.A., Evans C.A., Gocayne J.D.,				
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,				

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Vandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Mklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Durdon K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durkin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Matei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merulov G., Milshina N.V., Modary C., Morris J., Moshirefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclebo J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Relbert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Sieden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spirdling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitslks R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan W., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 [2]
 RP SEQUENCE FROM N.A.
 RA Celinker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
 RA Banson J., An H., Baldwin D., Banson J., Beeson K.Y., Busam D.A.,
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorese V., Doup L.E., Doyle C., Dreene D., Farfan D.,
 RA Ferrera S., Fries E., Galle R.F., Garg N.S., George R.A.,
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Ibegwan C., Jalali M., Kruse D., Li P., Matei B., Moshirefi A.,
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
 RA Paclebo J., Paragas V., Park S., Patel S., Pfeiffer B.,
 RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
 RA Stapleton M., Strong R., Svitslks R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
 RT "Sequencing of *Drosophila melanogaster* genome.";
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 [3]
 RP SEQUENCE FROM N.A.
 RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
 RA Hradecky P., Huang Y., Kaminker J.S., Prochick S.E., Smith C.D.,
 RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celinker S.E.,
 RA Clamp M., Drysdale R., Emmert D., Fries B., de Grey A., Harris N.,
 RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,
 RA Seale S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.,
 RT "Annotation of *Drosophila melanogaster* genome.";
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 [4]
 RP SEQUENCE FROM N.A.
 RA Adams M.D., Celinker S.E., Gibbs R.A., Rubin G.M., Venter J.C.;
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 [5]
 RP SEQUENCE FROM N.A.
 RA Flybase;
 RL Submitted (SEP-2002) to the EMBL/Genbank/DBJ databases.
 [6]

RP SEQUENCE FROM N.A.
 RC STRAIN-Berkeley;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorese V., Farfan D., Fries E., George R.,
 RA Gonzalez M., Guanin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nunoo J., Paclebo J., Paragas V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celinker S.;
 RL Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AE003459; AAF48473.2; -
 DR EMBL; AY061353; AAL28901.1; -
 DR FlyBase; FBgn0030655; CG9213.
 DR InterPro; IPR006768; CwFJ_C1.
 DR InterPro; IPR006767; CwFJ_C2.
 DR Pfam; PF04677; CwFJ_N1; 1.
 DR Pfam; PF04676; CwFJ_N2; 1.
 SQ SEQUENCE 687 AA; 78148 MW; 7376BAB9CE1EE876 CRC64;
 Query Match 4.9%; Score 133; DB 5; Length 687;
 Best Local Similarity 22.1%; Pred. No. 0.011;
 Matches 66; Conservative 43; Mismatches 134; Indels 56; Gaps 12;
 QY 181 VFEDDERSTDSQQCSSEDEDEFEETAQVSPRGEKRRQRRAMPPLISVQVSNADWV 240
 DB 4 IQFSGEKDXARQELREAREAMIQAKERALERGREOREYELR-----GEADWML 54
 QY 241 --LYKRLHLKLMELCNVYQWHLLENCEMEPPPIFKGDPPFILPSFQSESTPTGFGSG 298
 DB 55 PALAKLEKPAKKRN-VSHKSRSSKSSKSKSRKRNRN-----SSSESTSSSSSF-- 107
 QY 299 KETSEDDRSQSRHMGESLSLKAGCD--LLLP--SPRYEKDPSRKKEWENAGN 352
 DB 108 ----SEDEKERRKKRSKRKESASEDEVEAPPLADNVYKPPQDDM----- 157
 QY 353 KIYMAADKITSKLTETKYKRKQOHLSAPKEVAV--EKGEPLGP---RCQDSPLL 405
 DB 158 ----MTSESLILKTFSEKREKPAKNEKA--QQIDAVPAKSGRELNPYKSNCTGPGF 211
 QY 406 ORPQLMDQGMRSPFSGPELLRQDKRP-----RSGSTSSLSVSRDADAQIO 455
 DB 212 QKPDDEERQAKPHSSSSAAGSSRGWRKPKAKAPSPRRSRKSKATSSDEEEEAQV 270
 RESULT 6
 ID Q95R63 PRELIMINARY; PRT; 751 AA.
 AC Q95R63;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE SD04745P.
 GN CG12750.
 OS *Drosophila melanogaster* (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; *Drosophila*.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorese V., Farfan D., Fries E., George R.,
 RA Gonzalez M., Guanin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nunoo J., Paclebo J., Paragas V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celinker S.;
 RL Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AY061596; AAL29144.1; -
 DR FlyBase; FBgn0032678; CG12750.
 DR InterPro; IPR003891; IF_eif4G_MA3.
 DR Pfam; PF02847; MA3; 1.
 DR SMART; SM00544; MA3; 1.
 SQ SEQUENCE 751 AA; 86453 MW; DEA5672C82EE71E7 CRC64;
 Query Match 4.8%; Score 130.5; DB 5; Length 751;
 Best Local Similarity 22.1%; Pred. No. 0.021;


```

Db      860 KODVLVESTAGLPKDNPRNTR-----FSINFTISIGLGTDLDRF----- 902
Qy      256 YIQMLDLNCEMBEPIFK-----GDP-----FILPSFOSSTPTGGRS 297
Db      903 -----LKNNAKSPVAIINAEILANAGNPFDRGAPAGNTKVAPSSSSSS-----S 948
Qy      298 GKEPSEDRSGSRHMGESLSLKAGGDLPLPSPKYKKPNSK-----KEMWENAG 351
Db      949 SSDTSED--SSEBSSSSSSSSSSSSSEPKKKRRKKDKKKSKKATKESKTK 1005
Qy      352 NKIYMAADKTIISKLTETKKRQOHNLSAPPEVVEKKGEPILPGQDSEPLQPHL 411
Db      1006 NKKKKKKKKKK-----EKEKEKQKRSKKKEKDKKKKKKKKAKKSKRRKKSQS 1058
Qy      412 MDQGMRSFSAGPELLRODKRPRSGTSSLSVSVRAEAQ 453
Db      1059 SD-----SSGSE--DSDKSTSSSDSSSSSDSDAEPQ 1090

```

RESULT 8

```

ID      017596      PRELIMINARY;      PRT; 2150 AA.
AC      017596;
DT      01-JAN-1998 (TREMBlrel. 05, Created)
DT      01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT      01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE      C25D7.3 protein.
GN      C25D7.3.
OS      Caenorhabditis elegans.
OC      Eukaryota; Metazoa; Nematozoa; Chromadorea; Rhabditida; Rhabditoidea;
OC      Rhabditidae; Peloderinae; Caenorhabditis.
OX      NCBI_Taxid=6239;
RN      [1]
RN      [2]
RP      SEQUENCE FROM N.A.
RA      Ainecough R.;
RL      Submitted (Oct-1996) to the EMBL/GenBank/DBJ databases.
RX      MEDLINE=99069613; PubMed=9851916;
RA      none;
RT      "Genome sequence of the nematode C.elegans: A platform for
RT      investigating biology.";
RL      Science 282:2012-2018 (1998).
DR      EMBL; Z81039; CAB02774.1; -.
DR      Wormpep; C25D7.3; CE08389.
DR      InterPro; IPR007087; Znf_C2H2.
DR      PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
KW      Metal-binding; zinc; zinc-finger.
SQ      SEQUENCE 2150 AA; 249768 MW; E6A29F205124B569 CRC64;

```

```

Query Match      4.6%; Score 124; DB 5; Length 2150;
Best Local Similarity 20.6%; Pred. No. 0.36;
Matches 114; Conservative 63; Mismatches 166; Indels 210; Gaps 24;

Qy      33 IIEELPDERPNHTKXSVSFREIVSL-----SHQVLLQ---NLVYILLEEFVK 79
Db      1177 IERLSEDDQLGDQSESDISFEIIVDLLSEGEVAVQVQNSVTSSSTSFESLVEDDEE 1236
Qy      80 GP-----SPGEKTIQ-VPEAKLAGFLRYISMONTAVIFDLDLDYRTAREPDTSPGLKC 133
Db      1237 HPEQLPVASSEKANNQIVPEVEGVSVPVTNQ-----EENVTSQG--P 1279
Qy      134 LKKVSGIGGAGANLYRQASMSFNIFYHALVCAVLTNQ-----ETTAAQ----- 177
Db      1280 TLQESSISPSSSHIV-----TVDELITGESGPRAETTPVAEBSPPKKSGKT 1326
Qy      178 -----YKQVLFEDDERSTDSOOCSSDEDIFETLQVSPRCKEKQWARMPL--L 228
Db      1327 TRGRPKYKAKENIKKIQRPRGCKEKAHEEP--VVEQOEVEPEVGEVHEVPAPAAQL 1385
Qy      229 SVQPVSNADWVWLIVRLHLKCMELCNNTYIQMLDLNCEMBE-----PIFKGDPFIIIPS 283

```

```

Db      1386 ETEPI-----EQIEEPDKVPEPIIEA-----LPL 1410
Qy      284 FQSSSTPTSTGCGSEKTPSED-----RSQSRHMG 315
Db      1411 FETSPVPAPEGNINPSRAHSSDDVQVVISSETDPNGINLVEQVONDKLTAVQYSTEEILG 1470
Qy      316 ESLSLKAGCG-----DLLPSPKYKKDPKSKKEMWEN--AGNKIYMAA 359
Db      1471 EYGEIDEGAPSPSEIYVHDEVLQDEVLQPNPKSKKKRRKKKTTPPHIAKARKVFT--- 1527
Qy      360 DKTIISKLTET-----YKRRQOHNLSA-----FPKEVVEKKGEP-----LQPR 398
Db      1528 --SISK--TEIELAPPTQOSRKRMANVSSEATATRRQKRAVEPNDSVGRVLTP 1583
Qy      399 GQDSEPLQPHLMDQGMRSFSAGPELLRODKRPRSGTSSLSV----- 446
Db      1584 PEDLHETERPGHVEEGEPSTLRTGR-----STASSVTSRSKRLFLSKNPP 1632
Qy      447 VRDAEQIQAWTN 459
Db      1633 VPRMRIOQAGTN 1645

```

RESULT 9

```

ID      0961X4      PRELIMINARY;      PRT; 809 AA.
AC      0961X4;
DT      01-DEC-2001 (TREMBlrel. 19, Created)
DT      01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT      01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE      GH01088P (CG10732-PB).
GN      CG10732.
OS      Drosophila melanogaster (Fruit fly).
OC      Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC      Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC      Ephydroidea; Drosophilidae; Drosophila.
OX      NCBI_Taxid=7227;
RN      [1]
RN      [2]
RP      SEQUENCE FROM N.A.
RA      STRAIN=Berkelley.
RC      Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA      Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA      Gonzalez M., Guarini H., Li P., Liao G., Miranda A., Mungall C.J.,
RA      Munoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA      Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL      Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN      [2]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=20196006; PubMed=10731132;
RA      Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA      Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA      George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA      Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA      Brandon R.C., Rogers Y.H., Blazer R.G., Champe M., Pfeiffer B.D.,
RA      Wan K.H., Doyle C., Baxter B.G., Helt G., Nelson C.R., Gaber G.L.,
RA      Abtli J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA      Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA      Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolintsov S.,
RA      Botkova D., Botchan M.R., Bouck P., Brokstein P., Brothier P.,
RA      Butcher J.C., Buesam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA      Churruarain J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA      de la Piedad B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA      Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunov B.C., Dunn P.,
RA      Dudin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,
RA      Foeller C., Gabrielista C.C., Gary N.S., Gelbart W.M., Glasser K.,
RA      Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA      Harris N.L., Harvey D., Heitman T.J., Hernandez J.R., Houck J.,
RA      Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibeagwam C.,
RA      Jalaal M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA      Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
RA      Laevo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA      Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA      Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

```

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusseken D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
RA Williams S.M., Woodger, Morley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang X., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of *Drosophila melanogaster*."
RL Science 287:2185-2195(2000).
RN [3]
RP SEQUENCE FROM N.A.
RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Bantz J., An H., Baldwin D., Banton J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Fartan D.,
RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.,
RT "Sequencing of *Drosophila melanogaster* genome."
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Miura S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hrddecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA Ruby J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA Clump M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Kronmiller B., Marshall J.B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutnick F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.,
RT "Annotation of *Drosophila melanogaster* genome."
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter J.C.,
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL: AY047499; AAK77231.1;
DR EMBL: AE003538; AAN11838.1;
DR FlyBase: FBgn0036365; CG10732.
DR InterPro: IPR005613; AIP3.
DR Pfam: PF03915; AIP3; 1.
SQ SEQUENCE 809 AA; 89641 MW; 9862841AIF5FA81 CRC64;

Query Match 4.5%; Score 122; DB 5; Length 809;
Best Local Similarity 21.1%; Pred. No. 0.13;
Matches 105; Conservative 72; Mismatches 181; Indels 140; Gaps 23;

13 DTGSPKTPNNFHAQSCQIILIELP-PDEKPNHGHTKSVSFREIYVLSLHQVLLQNLQYD 71
23 DVALSEMSDNRAQSIQSLHVSLEFPTPTTP-----STDELQQRLEAANRNMQHLD 74
72 -----ILIEFVVK-----GSPGSEBKTIQVE--AKLAGFLR- 101
75 EQAKLIQIQNLAKTINEMERLRQQAQSLPHNVNGEAPKTESVOQVDDMAALVGRMKN 134
102 ---VISMQN-LAVIF-----DLT-----LQSYRAREPDISPGKCLKLVKVGIGC 143
135 LTAFLHNEHLSVYLGDDGPETLAEQALQKXESLKTQRE-----DWRNLVDLINSINR 189

QY 144 AANLYRQSAFNIYFHALVCAVLTNQETITABQVKVLFEDDERSTDSSQCCSDEDEDI 203
DB 190 AA---RETA-----RVIKKEKETPTPPK----- 211
QY 204 FEETAOVSPPRGKEKQWRAMPULSVQVPSNADWWLVKRLKLCMELCNNTYIQHLDL 263
DB 212 -EEAAPPAERXRVVPEYQNVPIIR-QEAAAN-----AQAALAAQA-----INQKTADI 260
QY 264 ENCEEPEPIKGDPPFLPSFOSESSPTSGSGKTEPSEDDRSQREHMGSSLSIKAG 323
DB 261 EALKQAQARKLG---MLYTSQIBESTPSMG--STERNSEERTSYRELPAETIAQRVFA 315
QY 324 GCDLLPSPSPVKPKOPSRKKEWENAGNKITYMAADKTIKSLMTEYKKRQOHNLGSAFP 383
DB 316 LNDV---TSELAAABAASLQKE-----RDRILA-LKAEIERKQQAALAAVQM 357
QY 384 KEVYVERKGGPPLPGRGDSFLRPPQHLMOGQNRHSFSAQ--ELLRQKRPRSGTGS 441
DB 358 GDDALKRSSLTPTP---TMRQRMTEEBDPSEVDNLSQATPTKEQLRDELRSQCERLRT 413
QY 442 SLVSVDAAEQIQAWTN 459
DB 414 EYEQKRELEQRYVASNN 431

RESULT 10
QY0V76 PRELIMINARY; PRT; 1657 AA.
AC QY0V76
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DT 01-OCT-2002 (TREMblrel. 22, Last annotation update)
DE CG10732 protein.
GN CG10732
OS *Drosophila melanogaster* (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; *Drosophila*.
OX NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STEIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolintsov S.,
RA Botkova D., Botchan M.R., Bouck J., Brokskein P., Brottier P.,
RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fowler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glöck A., Gong F., Gottrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houson K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mervinov G., Milshina N.V., Mobery C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusseken D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195 (2000).
 DR EMBL: AB003538; AAF49814.2; -;
 DR Flybase: FBgn0036365; CG10732.
 SQ SEQUENCE 1657 AA; 183992 MW; FFF5F036A233ECBD CRC64;
 Query Match 4.5%; Score 122; DB 5; Length 1657;
 Best Local Similarity 21.1%; Pred. No. 0.37; Mismatches 181; Indels 140; Gaps 23;
 Matches 105; Conservative 72;
 QY 13 DTQSPKTPNNFDHAQSCOLIIELP-PDEKPNHTKSVSPREIVSLSHQVLLQNTYD 71
 DB 301 DVALSEMSDNRAQSIQSLHSLVETPTPTP-----SFDELQQRLEASRNMQHLMD 352
 QY 72 ----LLEEFVK-----GPSGEEKTQVPE-AKLAGPLR- 101
 DB 353 EDAKLLQIOMLAKTHLEMERLROQADSLPHNVNGGEAPKYESVOQVDMAQLVGRMKN 412
 QY 102 ---YISMON-LAVIF-----DLL-----LDSVTRAREPDTSGUKLKYSGIGG 143
 DB 413 LTAFTHNQELSTVAGDGPETLAEQALQEKLESILRQRE-----DMKNLVDELNSINR 467
 QY 144 AANLYRQASMFNIYFHALVCAVLINQETITAEQVKVLFEDDERSTDSQCSSEDEDI 203
 DB 468 AA---RETA-----RVIKKEKEPTPTPPPK----- 489
 QY 204 FEETQVSPPRKPKRQMRAMPPLSVQVSNADVWLKRLKLCMELCNNTYQMHLDL 263
 DB 490 -EEAAPAFKERVVPEYQNRVPIIR-QEANA-----AQRALHAQAM-----INOKTADI 538
 QY 264 ENCMEEPTFKGDPFPLPSFQSESTPTSGFGSGKEPTSEDDRSQRHMGESLSLQAG 323
 DB 539 EALKQOMARKG---MLYVSOIEESTPSMG--STLERSEERTSYERLPAELIQRYFA 593
 QY 324 GGDLLLPSPKVKEDPSRKKEWENAGNKIYTMADKTIISKLMTEYKKRKOHNLSAFP 383
 DB 594 INDV-----TSELRAEALSLQKE-----RDRLLA-LKAEIERRQQAQAAVAQM 635
 QY 384 KEVKYKKEGEPGPGQDSPLIQROHLMDOQMHSFSAQ--ELLRODKRPRGSGTGS 441
 DB 636 GDDALKRSSLRPTP---TPMRQMTTEEDPSEVDNSIQATPTKSQLRDELRSQCERLUT 691
 QY 442 SLSSVSRDAEAQIOAWTN 459
 DB 692 EYEQKRELEQRYVASNN 709
 RESULT 11
 Q8FT17 PRELIMINARY; PRT; 1433 AA.
 AC Q8FT17;
 DT 01-JUN-2002 (TREMblrel. 21, Created)
 DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
 DT 01-OCT-2002 (TREMblrel. 22, Last annotation update)
 DE POL protein (Fragment).
 CN POL.
 OS *Saccharomyces bayanus* (Yeast) (*Saccharomyces uvarum*).
 OC Eukaryota; Fungi; Ascomycota; *Saccharomycotina*; *Saccharomycetes*;
 OC *Saccharomycetales*; *Saccharomycetaceae*; *Saccharomyces*.
 OK NCBI_TaxId=4931;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CLIB53; TRANSPOSON=LTR-retrotransposon Tsu4;
 RT Neugebisse C., Feldmann H., Bon E., Galliardin C., Casaregola S.;
 RA "Genomic evolution of the LTR-retrotransposons in hemiascomycetous
 RT yeasts";
 RL Submitted (MAR-2002) to the EMBL/Genbank/DBJ databases.

DR EMBL: AJ439550; CAD29535.1; -;
 DR InterPro: IPR001584; Rve.
 DR Pfam: PF00665; Irv; 1.
 FT NON TER
 SQ SEQUENCE 1433 AA; 164312 MW; 87C4075851F039BA CRC64;
 Query Match 4.3%; Score 116.5; DB 3; Length 1433;
 Best Local Similarity 19.9%; Pred. No. 0.91; Mismatches 142; Indels 107; Gaps 16;
 Matches 78; Conservative 65;
 QY 167 LTNQETITAEQVKVLFEDDERSTDSQCSSEDEDIFEETAQVSPRGKPKRQMRAMP 226
 DB 584 VTNDVDVSELQSKNITTESR--TNEINKPNTDEVDYEEVRYLPTALQEN----- 633
 QY 227 LLSVQ---VSNADVWLKRLKLCMELCNNTYQMHLDLENCEBEPPIFKGDPFPLPS 283
 DB 634 LVGQKQKTINNEENI--ASRMQKNI SGNEINYELELSDSDSCS----- 675
 QY 284 FQSSSTPTSGFGGKETPSEDDRSQRHMGESLSLQAGGDLPLPSPKVEKKDPSRK 343
 DB 676 -LHDSYDSTVITTSKQDLTDKQLOQSQELFEKVS-----DPEVLBEHMKIERQVESQN 729
 QY 344 KEMMENAG-----NKIYTMADKTIISKLMTEYKKRKOHNL-----SAFPKPKV 387
 DB 730 SDNETSQVQRPESIVTSSSDNDTON--DDYSTDKESHLLPLVYVNMNDTQTYDKPKN 787
 QY 388 VEKKEG---PLGPGQDSPLIQROHLMDO----- 414
 DB 788 EKSNNNSDISISPGNNEELVQ---LVDSNKAKEQATLESAITDEPIENPAANKA 843
 QY 415 GQMHSPAGPELLRODKRPRSGSTGSSLSVSVDAERQIAQWMMULTYVINOQIIPDQ 474
 DB 844 GFLKKAEPSS---LKKRRRPLEKNT-SFNDTAKKDNKRQRK-----NIIKLIPDN 889
 QY 475 TETALQPAVFPCC-ISQLTCHVTDIRVQAVRE 505
 DB 890 TETTSAPRIKITYNEAISRNADLKEKHAYKE 921
 RESULT 12
 Q8CD17 PRELIMINARY; PRT; 1110 AA.
 AC Q8CD17;
 DT 01-MAR-2003 (TREMblrel. 23, Created)
 DT 01-MAR-2003 (TREMblrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
 DE Hypothetical protein.
 CN Mus musculus (Mouse).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OK NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RT MEDLINE=22354683; PubMed=12466851;
 RA The PANTOM Consortium,
 RA The RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs";
 RL Nature 420:563-573 (2002).
 DR EMBL: AK029994; BAC6722.1; -;
 KM Hypothetical protein.
 SQ SEQUENCE 1110 AA; 128516 MW; 25406883552BBEB CRC64;
 Query Match 4.3%; Score 116; DB 11; Length 1110;
 Best Local Similarity 18.8%; Pred. No. 0.69; Mismatches 227; Indels 142; Gaps 25;
 Matches 107; Conservative 94;
 QY 4 AMAQVFWL---DTQSPKTPNNFDHAQSCOLIIEPPD---EKPNHTKSVSPREIV 56
 DB 60 SLRDLLIMLGGDRGQLETPNYLEDPAISQVAISPINEVLCPPSP-GNLMRNEFLVNR 118
 QY 57 VSLSLH--QVLLQNLVYLLE-----FVKGSPSGEERTTQVPAKLAGF 99


```

Db 119 MCVLENIQLQSKMNI PHLOTEKESNPQKTAFLNDQNLICQEHSGKLLQLLEVWNLROQ 178
Qy 100 LRYI-----SMQNLAVIFDILLDSYRTA-----REFPS-----128
Db 179 LKVAKEEDRAODEMORLTATLETETKKAANAIVIEELKTKRKMKIKIQLERROLAQE 238
Qy 129 -----PGKCLLKRVSGIGGAANLYROSAMFNIFYHALVCAVLINQETITAEQVKK 180
Db 239 KLVRESLETSSAMLLKQVKGWVVEARQVYHLOQECTSLHSMKTTQDILLQEQRK- 297
Qy 181 VLFEDDERSTDSQQCSSEDED-----IFEE--TAQVSPRKGKRRQMRAPMLLSVQPV 233
Db 298 ---MEDLCMTISQKSLDLSNRDNLICKLVENKATQISLKKEHENIY-LKSEILSLQDV 353
Qy 234 SNAWMLVYKRLHLCMELCNVYIQMHLDLEN-----CMEEPRIFKGDPPFLPSQSES-- 288
Db 354 SGKAQV-LNDQLSKSCSEL-TSMLOV-VKMENSRIIAHQAILKVEQCKITETQEOQWL 410
Qy 289 STPTGSGFSGETPSSEDDRSQREHMGESLTKAGGDLPLPSPKVEKQPSRKEKWE 348
Db 411 LDAHAASITGELQAVNQNKALQIHLDH-----LLEHNQCLQSGEAEKR----- 456
Qy 349 NAGNKIYMAADKTIISKLMTEYKRRKQ-----QNLSAFPKEVKEK----- 390
Db 457 ---TVQKELLESTIARLOQELKSLQEKSLLEKNWFOREINTEKEVAKENLEKE 513
Qy 391 ---KGEPLGPRGQD-SPLQRPQHLMD-----QGMKHSFSGAPPELLRODK 432
Db 514 LAESKEDLVNLQNLQTLMEENKHLTNKMASLEHHKATSDYQGVKEXAL-----EKITDSK 569
Qy 433 RPRSGTSSLSVSVRAEAQIOAMTNVNL 462
Db 570 NMLAYEKG-KIQTKVQKLEAQHLTPAETML 598

```

```

RESULT 13
Q8WZV3 PRELIMINARY; PRT; 2309 AA.
AC Q8WZV3;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Conserved hypothetical protein.
GN B7N14.260.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Aign V., Hohelsel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.,
RA Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RA Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL669986; CAD21060.1; -.
DR InterPro; IPR001611; LRR.
DR Pfam; PF00560; LRR; 6.
KW Hypothetical protein.
SQ SEQUENCE 2309 AA; 252994 MW; 583ACFED3A12B67C CRC64;

```

```

Query Match 4.3%; Score 115; DB 3; Length 2309;
Best Local Similarity 18.8%; Pred. No. 2.5;
Matches 95; Conservative 71; Mismatches 174; Indels 164; Gaps 23;
Qy 66 LQNLVDLLEFVKGSPGEEKTIQVPAKLAGF--LRYSIMQNLAVIFDILLDSYRTAR 123
Db 491 LNDISDLPEESLQSTPK-----RFSVKTEIEMDDYEHVPAQVSGHISDEI--SPKTRS 543
Qy 124 EFDTSFGKCLLKRVKVS-----GIGGANLYROSAMFNIFYHALVCAVLINQETITAEQV 179

```

```

Db 544 AVPRIPULSVSTRKRVASSQCGAGSPLKFPQ-----YDTFTSEILQ 584
Qy 180 KVL--FEDDE--RSTDSQQCSSEDE---DIFETAVQSP-----RGKEKQMRAMPL 227
Db 585 RRLSQYHDDTPGEOASSGQDTGDESPVHVSVMYSGGLSPKTHRRGQETGRKASVAV 644
Qy 228 -----LSVQVPSNADWMLVKRLHKLQMLCNVYIQMHLDLEN----- 265
Db 645 NQAGAVLDIGERODDPSTYSVGAAG-----IEBQKENGQND 683
Qy 266 CMEEPR-PIFGDPPFLPSQSESSTPTGFSG-----KETPSEDRSQSEHMGES 317
Db 684 SQEPRIPVPD-----FAPPSRSBADGDADANSNSVRKKQKPADASASKHSRWGSS 739
Qy 318 LSLKAGGDLPLP-----PSPKVEK-----KPSKKE-----WEN 349
Db 740 MAHKRHAFTHTIPATFTGTPSKRSDPAPEKRRPTSPSKDPTPKRRRTLESIDIEFGLD 799
Qy 350 AGNKIYMAADKTIISKLMTEYKRRKQOHNLSAFPKEVKEK---KGEPLGPRGQDPLQ 406
Db 800 ISPSPEFSKMLSYQQLQAIEKRR-----KEYNAEKHEQVPEPTPSGKDVPRR 850
Qy 407 RQHLMDQGMKHSFSGAPPELLRODKRPRS-----GSTGSSLSV-- 445
Db 851 TPTFLQSSSQKQCKGRPPLTELITKPSRPTPRSSKASGKRTPKHGGPGSAMNTDR 910
Qy 446 --SVRAEAQIOAMTNVLTVLNQ 467
Db 911 KPSIKTEDFLIEA--NKIMAMIRK 932

```

```

RESULT 14
Q8TE71 PRELIMINARY; PRT; 1077 AA.
AC Q8TE71;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE BBG1L.
GN BBG1L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Aerbaajinaï W., Miller J.L.,
RA Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY074490; ALU71549.1; -.
DR InterPro; IPR001073; C1q.
DR Pfam; PF00386; C1q; 1.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
SQ SEQUENCE 1077 AA; 120974 MW; 2888BF3C47D032D6 CRC64;

```

```

Query Match 4.2%; Score 114; DB 4; Length 1077;
Best Local Similarity 18.0%; Pred. No. 0.99; Indels 170; Gaps 21;
Matches 106; Conservative 91; Mismatches 223;
Qy 22 NFPDHAQSCQ-----LIEELPPDEKP---NGHTKKSVPFRIYVLSLHSHQVLLQNLVDIL 73
Db 172 HNLFEAELOKTSRGLSDLLKQKQKQREHMLKLEAEKKUKRTIIQVQVYLQNLQEH 231
Qy 74 LEEFVKGSPGEEKTIQVPAKLAGFLRYI-----SMQNLAVIFDILLDS 118
Db 232 VQXDFKG---GLNGAVLVLPSEKELDYLIKFSKLTCPERNESLSVEDQMEOSLSYFWDLLEG 288
Qy 119 YRTAREEDTSPGKCLLKRVSGIGGAANLYROSAMFNIFYHALVCAVLINQETITAEQV 178
Db 289 SEKAVGVTYKHLKDLISKLNSG-----YFESI PVPKNAKEKEVPLEE- 332
Qy 179 KVLVFEDEDERSTDSQQCSSEDEDIFE-ETAVQVSP----- 212

```

```

Db      333  EMLIQSEKKTOLSTESVSESESLMEFPQPEIQOPEILNRRYTBVDYSNKQEEQPE 391
      213  -----PRGKEKRO-----WRA-----RMPLLSVQPVSNADWMLV 242
Qy      332  ADVARKPNLPRKMDLTPEDGQKQESFKSWEAGKQEVSKPAVSLE----- 440
Db      243  KRLHLCEMLCNNTYOMHLIDLENCEBEPPIFKGDPFLLPSFQSSSTPSTGCF-----S 297
      441  -----QRKQDTSKLRLSTLPEBQKQOISKSKP-----SPSQWKQDTPSKXAGYQVEBK 489
Qy      298  GKEFP-----SEDDRSQSRHMGESL-SLKAGGDLPLPSPKYEKKQPSRKKEW 347
      490  KOETKRLNPFVQIQKQODPKKQTPKSWTPSMQSEQWTTTSWTTMCEBQDSKQEPETPKSWE 549
Db      348  ENAGKRIYTMADKTIISKLTMTYKKRKQOHNLSAPKEVYKKEEPL---GPRQDSPL 404
      550  NVVESQKHSILTSQSIQSPKSGVATASLIPNDQLPRKLTNTEPKDVPKPVHQPVGSSSTL 609
Qy      405  LQRP-----QHLMDQGMHSFSAGPELLRQDK-----RPRSGTSSLSVSVD 449
      610  PDPVLRRKEKLODLMTOIQGTGNFQO-BEVLDPDKPSAIPTSQPPSATPGSPVASKQON 668
Db      450  AEAQ-----IQAWTNMVLTVLNOIQILPDOT-----FTALQPAVF 484
      669  LSSQSDFLQEPQLQATSSPVTCSNACLVTTDQASSGSETEFWTSETPEVF 718

```

```

RESULT 15
Q8J1P0
ID Q8J1P0 PRELIMINARY; PRT; 565 AA.
AC Q8J1P0:
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE SOX9 longer form.
GN SOX9.
OS Oryzias latipes (Medaka fish) (Japanese ricefish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Belontiiformes; Adrianchthyidae; Oryziinae; Oryzias.
OX NCBI_TaxID=8090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2219435; PubMed=12211055;
RA Yokoi H., Kobayashi T., Tanaka M., Nagahama Y., Wakamatsu Y.,
RA Takeda H., Arai K., Morohashi K., Ozato K.;
RT "sox9 in a teleost fish, medaka (Oryzias latipes): evidence for
RT diversified function of sox9 in gonad differentiation.";
RL Mol. Reprod. Dev. 63:5-16(2002).
DR EMBL, AB078924; BAC06353.1; -.
DR InterPro; IPR000910; HMG_12_box.
DR Pfam; PF00505; HMG_box; 1.
DR SMART; SM00398; HMG; 1.
SQ SEQUENCE 565 AA; 61788 MW; 107670D94B70F24A CRC64;

```

```

Query Match 4.2%; Score 112; DB 13; Length 565;
Best Local Similarity 19.3%; Pred. No. 0.58;
Matches 67; Conservative 48; Mismatches 112; Indels 120; Gaps 15;

```

```

Qy      204  FEETQVSPRGKEKROKARAPL-----LSVQPVSNADWMLTVLREL--HKLIC 249
      17  FRSRAELGFTGGLQAOQWLSIVLPRCRGTAGASQRRSFRSDSSSFIITISRVVWHRSA 76
Qy      250  --MELCNNTYOMHLIDLENCEBEPPIFKGDPFLLPSFQSSS--TPSTGFSGKETPS-- 303
      77  TRMNLDPYLVKTEBQDKLSLDA-----SPMSQSDSAGSPASPSGSDTENTRP 127
Db      304  -----EDDRSQS--REHMGESLSLKAGGDLILPSP-----K 334
      128  RENGIMRADGALSDPKQEDDKFPACIRAVSQVLK---GYDWTLVPMFVRVNGSTKNK 183

```

```

Qy      335  VEKQPSRKKEWENAGNK-----ITYMADKTIISKLTMTYKKRKQOHNLSAPKEV 386
      184  PHVRRPMNAFVWVMAQAAARRKLADQYPHLHNAELSKTIGKLMR----- 225
Qy      387  KVEKKGEPLGPRGQDSPILQRPQHLMDQGMHSFSAGPELLRQDKRPRSGTSSLSVS 446
      226  -----LNEGKERPPVEAEKRLRVQHKDH-----PDYKYQPRRKSVKSGS---E 269
Qy      447  VRDAEAOIQAWTNMVLTVLNOIQILPDOTFTALQPAVFPCISQLTCH 493
      270  AEDGGEHIS--TNAI-----FKALQADSPASSMGEVH 300

```

```

Search completed: July 24, 2003, 11:56:56
Job time : 120.28 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 12:16:28 ; Search time 159.268 Seconds
(without alignments)
839.285 Million cell updates/sec

Title: US-09-991-681-27

Perfect score: 518
Sequence: 1 RIRAMAQGVMLDQCSPKT.....VRQAVREMLGRVGRVYDIIV 518

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size : 0

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_prodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	518	100.0	1770	4 Q9ULH6	Q9ULH6 homo sapien
2	296	57.1	589	4 Q8N4Y4	Q8N4Y4 homo sapien
3	296	57.1	592	4 Q96CH9	Q96CH9 homo sapien
4	9	1.7	171	8 Q8HD83	Q8HD83 panulirus j
5	9	1.7	216	2 Q85197	Q85197 streptococc
6	9	1.7	216	16 Q97R75	Q97R75 streptococc
7	9	1.7	216	16 Q8DQ41	Q8DQ41 streptococc
8	9	1.5	121	4 Q9H7M8	Q9H7M8 homo sapien
9	8	1.5	328	2 Q8L254	Q8L254 proteus vul
10	8	1.5	426	11 Q8CI16	Q8CI16 mus musculu
11	8	1.5	426	11 Q8RTX3	Q8RTX3 mus musculu
12	8	1.5	427	10 Q9M035	Q9M035 arabidopsis
13	8	1.5	434	4 Q8IV13	Q8IV13 homo sapien
14	8	1.5	483	16 Q8Y7T3	Q8Y7T3 listeria mo
15	8	1.5	610	5 Q9XVNZ	Q9XVNZ caenorhabdi
16	8	1.5	621	16 Q9PFC6	Q9PFC6 xylella fas

17	8	1.5	1532	10 Q9SRD9	Q9SRD9 arabidopsis
18	7	1.4	33	10 Q945D4	Q945D4 castanea sa
19	7	1.4	67	10 Q8GVSO	Q8GVSO oryza sativ
20	7	1.4	76	2 Q8VRS6	Q8VRS6 escherichia
21	7	1.4	82	16 Q8X345	Q8X345 escherichia
22	7	1.4	86	16 Q922V9	Q922V9 rhizobium m
23	7	1.4	88	17 Q9HHM3	Q9HHM3 halobacteri
24	7	1.4	90	16 Q99XM6	Q99XM6 streptococc
25	7	1.4	113	10 Q8VXR5	Q8VXR5 cicor ariet
26	7	1.4	129	17 Q50122	Q50122 pyrococcus
27	7	1.4	137	16 Q9HU07	Q9HU07 pseudomonas
28	7	1.4	137	16 Q9R2T7	Q9R2T7 delnoccocus
29	7	1.4	141	16 Q9JWS0	Q9JWS0 neisseria m
30	7	1.4	142	5 Q8IMK8	Q8IMK8 dirosophila
31	7	1.4	143	5 Q15660	Q15660 plasmodium
32	7	1.4	143	10 P93272	P93272 malus domes
33	7	1.4	143	17 Q9YDK3	Q9YDK3 aetopyrum p
34	7	1.4	147	12 Q8UDX8	Q8UDX8 macaca fasc
35	7	1.4	149	16 P74153	P74153 synchocyst
36	7	1.4	161	16 Q92HT4	Q92HT4 rickettsia
37	7	1.4	164	5 Q16292	Q16292 caenorhabdi
38	7	1.4	164	16 Q8Z4U4	Q8Z4U4 salmonella
39	7	1.4	166	4 Q96NM6	Q96NM6 homo sapien
40	7	1.4	174	8 Q8WKP7	Q8WKP7 zephyra ele
41	7	1.4	174	8 Q8WKO0	Q8WKO0 walleria ma
42	7	1.4	174	8 Q8WKR5	Q8WKR5 lecopillaea
43	7	1.4	174	8 Q8WL48	Q8WL48 doryanthes
44	7	1.4	175	8 Q8WKS8	Q8WKS8 syringodea
45	7	1.4	176	8 Q8WKN8	Q8WKN8 chamanthe

ALIGNMENTS

RESULT 1

Q9ULH6 PRELIMINARY; PRT; 1770 AA.
AC Q9ULH6; Q96P46;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE BIG3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Hong W.;
RT "XIAI244 as a novel distantly related member (BIG3) of the BIG1/Sec7p
subfamily of ARP GTPs.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF413080; AL04174.1; -;
DR InterPro; IPR000904; Sec7.
DR SMART; SM00222; Sec7; 1.
KW Hypothetical protein.
SO SEQUENCE 1770 AA; 195845 MW; 5E996E36A6F92ABA CRC64;

Query Match 100.0%; Score 518; DB 4; Length 1770;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RIRAMAQGVMLDQCSPKTPNNPDHOSCOLIIEPPDEKPNQTKKVSFRIVSL 60
DB 1253 RIRAMAQGVMLDQCSPKTPNNPDHOSCOLIIEPPDEKPNQTKKVSFRIVSL 1312
QY 61 SHOVLQNTLYDILLIEFVKSPGEEKTIQVPEAKLAGFIRYISMOMAVIFDLILDSYR 120
DB 1313 SHOVLQNTLYDILLIEFVKSPGEEKTIQVPEAKLAGFIRYISMOMAVIFDLILDSYR 1372
QY 121 TAREPDTSPGLKCLLKVKVSGIGGANLYROSANSFNIFYFALVCAVUTNDETTAEQVK 180
DB 1373 TAREPDTSPGLKCLLKVKVSGIGGANLYROSANSFNIFYFALVCAVUTNDETTAEQVK 1432

QY 181 VLFEDDERSTSSQCCSEDEDEFEETAVSPRKEKRGWRAMPILLSVOPVSNADVM 240
Db 1433 VLFEDDERSTSSQCCSEDEDEFEETAVSPRKEKRGWRAMPILLSVOPVSNADVM 1492
QY 241 LVKRLHKLKCMELCNNTYIQMHLDLENCEMBEPIFKGDPEFILLPSFQSSSTSTGFSGKE 300
Db 1493 LVKRLHKLKCMELCNNTYIQMHLDLENCEMBEPIFKGDPEFILLPSFQSSSTSTGFSGKE 1552
QY 301 TPSEDDRSQSRHEHGESLSLAKAGGDLILLPSPKVEKDPBRKKEWMENAGNKITYMAAD 360
Db 1553 TPSEDDRSQSRHEHGESLSLAKAGGDLILLPSPKVEKDPBRKKEWMENAGNKITYMAAD 1612
QY 361 KTISKLMTEYKRRKQOHNLSAFPKEVKEKGEPLGRGQDSPLQRPQHLMDGQMRHS 420
Db 1613 KTISKLMTEYKRRKQOHNLSAFPKEVKEKGEPLGRGQDSPLQRPQHLMDGQMRHS 1672
QY 421 FSAGPELLRQDKRPRSGSTGSSLSVSRDAEAQIQAMTNMVLTVLNOIQILPDQTFALQ 480
Db 1673 FSAGPELLRQDKRPRSGSTGSSLSVSRDAEAQIQAMTNMVLTVLNOIQILPDQTFALQ 1732
QY 481 PAVPFCISQLTCHVTDIRVQAVREMLGRVGRVYDIIV 518
Db 1733 PAVPFCISQLTCHVTDIRVQAVREMLGRVGRVYDIIV 1770

RESULT 2

Q8N4Y4 PRELIMINARY; PRT; 589 AA.
AC Q8N4Y4;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Similar to KIAA1244 protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC033191; AAH33191.1; -.
FT NON TER 1
SQ SEQUENCE 589 AA; 66086 MW; 8041EEA348DE65F7 CRC64;

Query Match 57.1%; Score 296; DB 4; Length 589;
Best Local Similarity 99.6%; Pred. No. 2,4e-299;
Matches 516; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 RIRMAQOVFMDLTQCSPTKPNPFDAHSCQILIELPDEKPNHGTKKSVSFREIVSL 60
Db 73 RIRMAQOVFMDLTQCSPTKPNPFDAHSCQILIELPDEKPNHGTKKSVSFREIVSL 132
QY 61 SHQVLTQNLVYDILIEEFVKGPSPEEKTIOVPEAKLAGFLRYISQNLAVIFDILLDSYR 120
Db 133 SHQVLTQNLVYDILIEEFVKGPSPEEKTIOVPEAKLAGFLRYISQNLAVIFDILLDSYR 192
QY 121 TAREFDTPGKCLKLKAVSGIGGANLYROSAMSFNIFHALVCALVLTNOETTTAEQYKK 180
Db 193 TAREFDTPGKCLKLKAVSGIGGANLYROSAMSFNIFHALVCALVLTNOETTTAEQYKK 252
QY 181 VLFEDDERSTSSQCCSEDEDEFEETAVSPRKEKRGWRAMPILLSVOPVSNADVM 240
Db 253 VLFEDDERSTSSQCCSEDEDEFEETAVSPRKEKRGWRAMPILLSVOPVSNADVM 311
QY 241 LVKRLHKLKCMELCNNTYIQMHLDLENCEMBEPIFKGDPEFILLPSFQSSSTSTGFSGKE 300
Db 312 LVKRLHKLKCMELCNNTYIQMHLDLENCEMBEPIFKGDPEFILLPSFQSSSTSTGFSGKE 371
QY 301 TPSEDDRSQSRHEHGESLSLAKAGGDLILLPSPKVEKDPBRKKEWMENAGNKITYMAAD 360

Db 372 TPSEDDRSQSRHEHGESLSLAKAGGDLILLPSPKVEKDPBRKKEWMENAGNKITYMAAD 431
QY 361 KTISKLMTEYKRRKQOHNLSAFPKEVKEKGEPLGRGQDSPLQRPQHLMDGQMRHS 420
Db 432 KTISKLMTEYKRRKQOHNLSAFPKEVKEKGEPLGRGQDSPLQRPQHLMDGQMRHS 491
QY 421 FSAGPELLRQDKRPRSGSTGSSLSVSRDAEAQIQAMTNMVLTVLNOIQILPDQTFALQ 480
Db 492 FSAGPELLRQDKRPRSGSTGSSLSVSRDAEAQIQAMTNMVLTVLNOIQILPDQTFALQ 551
QY 481 PAVPFCISQLTCHVTDIRVQAVREMLGRVGRVYDIIV 518
Db 552 PAVPFCISQLTCHVTDIRVQAVREMLGRVGRVYDIIV 589

RESULT 3

O96CH9 PRELIMINARY; PRT; 592 AA.
AC O96CH9;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC014227; AAH14227.1; -.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 592 AA; 66400 MW; FA1E807B0DF47B5 CRC64;

Query Match 57.1%; Score 296; DB 4; Length 592;
Best Local Similarity 99.6%; Pred. No. 2,4e-299;
Matches 516; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 RIRMAQOVFMDLTQCSPTKPNPFDAHSCQILIELPDEKPNHGTKKSVSFREIVSL 60
Db 76 RIRMAQOVFMDLTQCSPTKPNPFDAHSCQILIELPDEKPNHGTKKSVSFREIVSL 135
QY 61 SHQVLTQNLVYDILIEEFVKGPSPEEKTIOVPEAKLAGFLRYISQNLAVIFDILLDSYR 120
Db 136 SHQVLTQNLVYDILIEEFVKGPSPEEKTIOVPEAKLAGFLRYISQNLAVIFDILLDSYR 195
QY 121 TAREFDTPGKCLKLKAVSGIGGANLYROSAMSFNIFHALVCALVLTNOETTTAEQYKK 180
Db 196 TAREFDTPGKCLKLKAVSGIGGANLYROSAMSFNIFHALVCALVLTNOETTTAEQYKK 255
QY 181 VLFEDDERSTSSQCCSEDEDEFEETAVSPRKEKRGWRAMPILLSVOPVSNADVM 240
Db 256 VLFEDDERSTSSQCCSEDEDEFEETAVSPRKEKRGWRAMPILLSVOPVSNADVM 314
QY 241 LVKRLHKLKCMELCNNTYIQMHLDLENCEMBEPIFKGDPEFILLPSFQSSSTSTGFSGKE 300
Db 315 LVKRLHKLKCMELCNNTYIQMHLDLENCEMBEPIFKGDPEFILLPSFQSSSTSTGFSGKE 374
QY 301 TPSEDDRSQSRHEHGESLSLAKAGGDLILLPSPKVEKDPBRKKEWMENAGNKITYMAAD 360
Db 375 TPSEDDRSQSRHEHGESLSLAKAGGDLILLPSPKVEKDPBRKKEWMENAGNKITYMAAD 434
QY 361 KTISKLMTEYKRRKQOHNLSAFPKEVKEKGEPLGRGQDSPLQRPQHLMDGQMRHS 420
Db 435 KTISKLMTEYKRRKQOHNLSAFPKEVKEKGEPLGRGQDSPLQRPQHLMDGQMRHS 494
QY 421 FSAGPELLRQDKRPRSGSTGSSLSVSRDAEAQIQAMTNMVLTVLNOIQILPDQTFALQ 480
Db 495 FSAGPELLRQDKRPRSGSTGSSLSVSRDAEAQIQAMTNMVLTVLNOIQILPDQTFALQ 554

Qy 481 PAVFPCISQLTCHVTDIRVQAVREWLGRVGRVYDIIV 518
 Db 555 PAVFPCISQLTCHVTDIRVQAVREWLGRVGRVYDIIV 592

RESULT 4

Q8HD83 PRELIMINARY; PRT; 171 AA.
 AC Q8HD83;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE NADH dehydrogenase subunit 6.
 GN ND6.
 OS Panulirus japonicus (Japanese spiny lobster).
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Palinura;
 OC Palinuroidea; Palinuridae; Panulirus.
 OC NCBI_TaxID=6736;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22226630; PubMed=12242015;
 RA Yamauchi M.M., Miya M.U., Nishida M.;
 RT "Complete mitochondrial DNA sequence of the Japanese spiny lobster,
 RT Panulirus japonicus (Crustacea: Decapoda).";
 RL Gene 295:89-96(2002).
 DR EMBL: AB071201; BAC16326.1; -
 KW Mitochondrion.
 SQ SEQUENCE 171 AA; 19041 MW; 64C76412B14450AA CRC64;

Query Match 1.7%; Score 9; DB 8; Length 171;
 Best Local Similarity 100.0%; Pred. No. 1.3;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 226 PLUSVQPV 234
 Db 103 PLUSVQPV 111

RESULT 5

O85197 PRELIMINARY; PRT; 216 AA.
 ID O85197;
 AC O85197;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Competence protein.
 GN CBLA.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OC NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Rx;
 RX MEDLINE=98241533; PubMed=9573156;
 RA Pestova E.V., Morrison D.A.;
 RT "Isolation and characterization of three Streptococcus pneumoniae
 RT transformation-specific loci by use of a lacZ reporter insertion
 RT vector.";
 RL J. Bacteriol. 180:2701-2710(1998).
 DR EMBL: AF052208; AAC23741.1; -
 DR InterPro: IPR004787; COME_-
 DR InterPro: IPR004509; COMEA_HHH.
 DR InterPro: IPR004445; HH.
 DR InterPro: IPR003583; HH_1.
 DR Pfam: PF00633; HH; 2.
 DR SMART: SM00278; HH; 2.
 DR TIGRFAMs: TIGR01259; COME. 1.
 DR TIGRFAMs: TIGR00426; TIGR00426; 1.
 RN SEQUENCE 216 AA; 23244 MW; C96DA1761B81938 CRC64;

Query Match 1.7%; Score 9; DB 2; Length 216;
 Best Local Similarity 100.0%; Pred. No. 1.6;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 135 LKRVSGIGG 143
 Db 195 LKRVSGIGG 203

RESULT 6

O97R75 PRELIMINARY; PRT; 216 AA.
 ID O97R75;
 AC O97R75;
 DT 01-OCT-2001 (TREMBlrel. 18, Created)
 DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Competence protein Cella.
 GN SP0954.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OC NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TIGR4;
 RX MEDLINE=21357209; PubMed=11463916;
 RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
 RA Peterson S., Heidelberg J., Deboy R.T., Haft D.H., Dodson R.J.,
 RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
 RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
 RA Holtzapple E., Khouli H., Wolf A.M., Uetzerback T.R., Hansen C.L.,
 RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
 RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
 RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser J.C.M.;
 RT "Complete genome sequence of a virulent isolate of Streptococcus
 RT pneumoniae.";
 RL Science 293:498-506(2001).
 DR EMBL: AE007400; AAK75075.1; -
 DR TIGR: SP0954; -
 DR InterPro: IPR004787; COME.
 DR InterPro: IPR004509; COMEA_HHH.
 DR InterPro: IPR004445; HH.
 DR InterPro: IPR003583; HH_1.
 DR Pfam: PF00633; HH; 2.
 DR SMART: SM00278; HH; 2.
 DR TIGRFAMs: TIGR01259; COME. 1.
 DR TIGRFAMs: TIGR00426; TIGR00426; 1.
 KW Complete proteome.
 SQ SEQUENCE 216 AA; 23181 MW; 291462B84608B4A CRC64;

Query Match 1.7%; Score 9; DB 16; Length 216;
 Best Local Similarity 100.0%; Pred. No. 1.6;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 135 LKRVSGIGG 143
 Db 195 LKRVSGIGG 203

RESULT 7

O8DQ41 PRELIMINARY; PRT; 216 AA.
 ID O8DQ41;
 AC O8DQ41;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Competence protein.
 GN CBLA OR SP0856.
 OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OC NCBI_TaxID=171101;
 RN [1]

```

RP  SEQUENCE FROM N.A.
RX  MEDLINE=21429245; PubMed=11544234;
RA  Hoskins J., Albom W.E. Jr., Arnold J., Blaszcak L.C., Burgette S.,
RA  Dehoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
RA  Gilmour R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,
RA  Leland D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
RA  McBreth S.M., McMeney M., McMaster K., Mundy C.W., Nicas T.I.,
RA  Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rockey P.,
RA  Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
RA  Zook C.A., Baltz R.H., Jaskunas S.R., Kosteck P.R. Jr., Skatrud P.L.,
RA  Glass J.I.;
RT  "Genome of the bacterium Streptococcus pneumoniae strain R6.";
RL  J. Bacteriol. 183:5709-5717(2001).
DR  EMBL; AF008460; AAK99660.1; -.
KW  Complete proteome.
SQ  SEQUENCE 216 AA; 23230 MW; C4FDBE7761B80638 CRC64;

Query Match 1.7%; Score 9; DB 16; Length 216;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 135 LKXVSGIG 143
Db 195 LKXVSGIG 203

RESULT 8
O9H7W8 PRELIMINARY; PRT; 121 AA.
AC O9H7W8;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Hypothetical protein FLJ14166.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isegai T., Oca T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Wagatsuma M., Hobeiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuno Y.,
RA Niimoriya K., Iwayanagi T.;
RT "NDO human cDNA sequencing project.";
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK024228; BAB14856.1; -.
DR EMBL; BC013353; AAH13353.1; -.
KW Hypothetical protein.
SQ SEQUENCE 121 AA; 12905 MW; 37ACFP89B40243C9 CRC64;

Query Match 1.5%; Score 8; DB 4; Length 121;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 436 SGTGSSL 443
Db 57 SGTGSSL 64

RESULT 9
O8L254 PRELIMINARY; PRT; 328 AA.
AC O8L254;
DT 01-OCT-2002 (TREMBlrel. 22, Created)

```

```

DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Plasmid partition protein.
GN PARB.
OS Proteus vulgaris.
OG Plasmid RstI.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Proteus.
OX NCBI_TaxId=585;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UR-75;
RA Murata T., Hayashi T.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=UR-75;
RA Murata T., Ohnishi M., Ara T., Kaneko J., Han C.-G., Li Y.F.,
RA Takashima K., Nojima H., Nakayama K., Kaji A., Kamio Y., Miki T.,
RA Mori H., Ohtsubo E., Terawaki Y., Hayashi T.;
RT "Complete Nucleotide Sequence of Plasmid RstI: Implications for
RT Evolution of Large Plasmid Genomes.";
RL J. Bacteriol. 184:3194-3202(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=UR-75;
RA MEDLINE=96184644; PubMed=8645296;
RA Tian Q.B., Ohnishi M., Tabuchi A., Terawaki Y.;
RT "A new Plasmid-encoded proteic killer gene system: cloning,
RT sequencing, and analyzing his locus of plasmid RstI.";
RL Biochem. Biophys. Res. Commun. 220:280-284(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=UR-75;
RA MEDLINE=9435883; PubMed=8076071;
RA Janosi L., Yonemitsu H., Hong H., Kaji A.;
RT "Molecular cloning and expression of a novel hydroxymethylcytosine-
RT specific restriction enzyme (pvrtsII) modulated by glucosylation of
RT DNA.";
RL J. Mol. Biol. 242:45-61(1994).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=UR-75;
RA MEDLINE=91193219; PubMed=2013575;
RA Mochida S., Tsuchiya H., Mori K., Kaji A.;
RT "Three short fragments of RstI DNA are responsible for the
RT temperature-sensitive growth phenotype (tes) of host bacteria.";
RL J. Bacteriol. 173:2600-2607(1991).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=UR-75;
RA MEDLINE=88289863; PubMed=2840681;
RA Nozue H., Tsuchiya K., Kamio Y.;
RT "Nucleotide sequence and copy control function of the extension of the
RT incI region (incI-b) of RstI.";
RL Plasmid 19:46-56(1988).
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN=UR-75;
RA MEDLINE=86139175; PubMed=3277947;
RA Tanaka M., Okawa N., Mori K., Suyama Y., Kaji A.;
RT "Nucleotide sequence of an RstI fragment causing temperature-dependent
RT instability.";
RL J. Bacteriol. 170:1175-1182(1988).
RN [8]
RP SEQUENCE FROM N.A.
RC STRAIN=UR-75;
RA MEDLINE=85234397; PubMed=2989253;
RA Mollet B., Clerget M., Meyer J., Iida S.;
RT "Organization of the Tn6-related kanamycin resistance transposon
RT Tn2680 carrying two copies of IS26 and an IS903 variant, IS903. B.";
RL J. Bacteriol. 163:55-60(1985).

```

RN [9]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UR-75;
 RX MEDLINE=84185439; PubMed=6325193;
 RA Kamio Y., Tabuchi A., Itoh Y., Katagiri H., Terawaki Y.,
 RT "Complete nucleotide sequence of mini-Rts1 and its copy mutant."
 RL J. Bacteriol. 158:307-312(1984).
 RN [10]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UR-75;
 RX MEDLINE=83290717; PubMed=6309744;
 RA Kamio Y., Terawaki Y.,
 RT "Nucleotide sequence of an incompatibility region of mini-Rts1 that
 contains five direct repeats."
 RL J. Bacteriol. 155:1185-1191(1983).
 RN [11]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UR-75;
 RX MEDLINE=68393387; PubMed=4876466;
 RA Terawaki Y., Kakizawa Y., Takayasu H., Yoshikawa M.,
 RT "Temperature sensitivity of cell growth in *Escherichia coli* associated
 with the temperature sensitive R(KM) factor."
 RL Nature 219:284-285(1968).
 DR EMBL; AF004237; BAB93856.1; -.
 DR InterPro; IPR003115; ParBc.
 DR InterPro; IPR004437; ParB_part.
 DR Pfam; PF02195; ParBc; 1.
 DR SMART; SM00470; ParB; 1.
 DR TIGRFAMs; TIGR00180; parB_part; 1.
 KM Plasmid.
 SQ SEQUENCE 328 AA; 37072 MW; 926B9924167A59P9 CRC64;

Query Match 1.5%; Score 8; DB 2; Length 328;
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 171 ETTAEQV 178
 Db 49 ETTAEQV 56

RESULT 10
 Q8CII6 PRELIMINARY; PRT; 426 AA.
 AC 08CII6;
 DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
 DE Similar to chromosome 20 open reading frame 9.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Kidney;
 RA Strauberg R.;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC037708; AAH37708.1; -.
 SQ SEQUENCE 426 AA; 48223 MW; EC0207BCB2327105 CRC64;

Query Match 1.5%; Score 8; DB 11; Length 426;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 478 ALQPAVFP 485
 Db 334 ALQPAVFP 341
 RESULT 11
 Q8BTX3 PRELIMINARY; PRT; 426 AA.
 ID Q8BTX3

AC Q8BTX3;
 DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE ND5 protein homolog.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NOD; TISSUE=Thymus;
 RX MEDLINE=22354683; PubMed=1246851;
 RA The PANTOM Consortium.
 RA "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs."
 RL Nature 420:563-573(2002).
 DR EMBL; AK088467; BAC40371.1; -.
 SQ SEQUENCE 426 AA; 48248 MW; EC1B2F1E21327105 CRC64;

Query Match 1.5%; Score 8; DB 11; Length 426;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 478 ALQPAVFP 485
 Db 334 ALQPAVFP 341

RESULT 12
 Q9M035 PRELIMINARY; PRT; 427 AA.
 AC Q9M035;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
 DE Hypothetical 46.1 kDa protein.
 GN T1008.80.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bevan M., Pohl T., Weizenegger T., Bancroft I., Mewes H.W., Rudd S.,
 RA Lemcke K., Mayer K.F.X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL161746; CAB81920.1; -.
 KM Hypothetical protein.
 SQ SEQUENCE 427 AA; 48124 MW; 4B89FAADBBD33B0F CRC64;

Query Match 1.5%; Score 8; DB 10; Length 427;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 49 SVSFREIV 56
 Db 145 SVSFREIV 152

RESULT 13
 Q8IV13 PRELIMINARY; PRT; 434 AA.
 AC Q8IV13;
 DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)

```

DE  Hypothetical protein FLJ14166.
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
OX  NCBI_TaxID=9606;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Brain;
RA  Strausberg R.;
RL  Submitted (JUL-2002) to the EMBL/Genbank/DBJ databases.
DR  EMBL; BC035871; AAH35871.1; -
KW  Hypothetical protein.
SQ  SEQUENCE 434 AA; 48274 MW; 01D44BEED823426D CRC64;

Query Match 1.5%; Score 8; DB 4; Length 434;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  436 SGTSSSL 443
    |||||
DB  370 SGTSSSL 377

RESULT 14
OBY7T3 PRELIMINARY; PRT; 483 AA.
AC  OBY7T3;
DT  01-MAR-2002 (TRENBLrel. 20, Created)
DT  01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT  01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE  Hypothetical protein lmo1188.
GN  LMO1188.
OS  Listeria monocytogenes.
OC  Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX  NCBI_TaxID=1639;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=EGDe / Serovar 1/2a;
RX  MEDLINE=21537279; PubMed=11679669;
RA  Glaeser P., Frangaul L., Buchrieser C., Rusniok C., Amend A.,
RA  Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA  Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA  Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA  Entian K.-D., Fshih H., Garcia-del Portillo F., Garrido P.,
RA  Gautier L., Goebel W., Gomez-Lopez N., Hain T., Haut U., Jackson D.,
RA  Jones L.-M., Kaerst U., Kretz J., Kunz M., Kunst F., Kurapkak G.,
RA  Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA  Nordstiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA  Remmel B., Rose M., Schueter T., Simoes N., Tierrez A.,
RA  Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT  "Comparative genomes of Listeria species.";
RL  Science 294:849-852(2001);
DR  EMBL; AL591978; CAC99266.1; -
KW  Listeria; LMO1188; -
SQ  SEQUENCE 483 AA; 56753 MW; 8092886531E2BDAS CRC64;

Query Match 1.5%; Score 8; DB 16; Length 483;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  173 ITAEQVK 180
    |||||
DB  390 ITAEQVK 397

RESULT 15
OQXVN2 PRELIMINARY; PRT; 610 AA.
AC  OQXVN2;
DT  01-NOV-1999 (TRENBLrel. 12, Created)
DT  01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT  01-MAR-2003 (TRENBLrel. 23, Last annotation update)

```

```

DE  F53B6.6 protein.
GN  F53B6.6.
OS  Caenorhabditis elegans.
OC  Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC  Rhabditidae; Peloderiinae; Caenorhabditis.
OX  NCBI_TaxID=6239;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  White S.;
RA  Submitted (OCT-1996) to the EMBL/Genbank/DBJ databases.
RN  [2]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=99069613; PubMed=9851916;
RA  none;
RT  "Genome sequence of the nematode C. elegans: A platform for
RT  investigating biology.";
RL  Science 282:2012-2018(1998).
DR  EMBL; Z81086; CAB03116.1; -
DR  WormPep; F53B6.6; CRI0900.
DR  InterPro; IPR001507; Endoglin/CD105.
DR  Pfam; PF00100; zona_pellucida; 1.
DR  SMART; SMO0241; ZP; 1.
SQ  SEQUENCE 610 AA; 68585 MW; 09819DC5C3741329 CRC64;

Query Match 1.5%; Score 8; DB 5; Length 610;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  286 SESSTPST 293
    |||||
DB  481 SESSTPST 488

```

Search completed: July 24, 2003, 12:25:43
 Job time : 163.268 secs